Some Topological Properties of Rough Sets with Tools for Data Mining

A. S. Salama
Department of Mathematics, Faculty of Science, Tanta University
Tanta, Egypt

Abstract
Rough set theory has a significant importance in many fields. For example, some branches of artificial intelligence, such as inductive reasoning, automatic classification, pattern recognition, learning algorithms, classification theory, cluster analysis, measurement theory and taxonomy. Also, in the domains of Medicine, Pharmacology, Banking, Market research and Engineering the rough set theory has demonstrated its usefulness. The main aim of this paper is to describe some topological properties of rough sets and open the door about more accurate topological measures of data mining.
Keywords: *Rough Sets; Topological Spaces; Rough Approximations; Data Reduction; Data Mining; Rule Extraction.*

1.Introduction

Rough set theory, proposed in [18-23], is a good mathematical tool for data representation. Its methodology is concerned with the classification and analysis of missing attribute values, uncertain or incomplete information systems and knowledge, and it is considered one of the first non-statistical approaches in data analysis [7,10,12].
The fundamental concept behind it is lower and upper approximations of a set, the approximation of sets being the formal tool of knowledge discovery and data mining [24, 25,30-33].
The subset generated by lower approximations is characterized by certain objects that will definitely form part of an interest subset, whereas the upper approximation is characterized by uncertain objects that will possibly form part of an interest subset. Every subset defined through upper and lower approximation is known as rough set. If we considered a topological space instead of approximation space we will generate topological rough sets using many topological notions such as pre-open sets[1,26-29,42].
Over the years rough sets have become a valuable tool in the resolution of various problems, such as: representation of missing (uncertain) or imprecise knowledge; knowledge analysis; identification and evaluation of date dependency; reasoning based an uncertain and reduct of information data [2-6,8,9,11].

The extent of rough sets applications used today are much wider than in the past, principally in the areas of data mining, medicine, analysis of database attributes and process control. The subject of this paper is to present the topological properties of rough sets [43-45].
The key to the present paper is provided by the exact mathematical formulation of the concept of approximate (rough) equality of sets in a given approximation space. An approximation space is understood as a pair (U, R), where U is a certain set called universe, and $R \subset U \times U$ is an indiscernibility relation. In the basic construction of the theory, R is assumed to be an equivalence relation.
The aim of this paper is to describe some topological properties of rough sets, introduced in [18] and investigated in [18-23].
Moreover, we give further study on Pawlak rough sets and introduce new examples and proofs to simplify some rough set theory concepts.
This paper is structured as follows:
In Section 2, we study the notion of rough sets in an approximation space and investigate some of its properties. The notion of topological rough sets and its relation with Pawlak rough sets are discussed in Section 3. The main goal of Section 4 is to spotlights on the notion of a topological rough classes and investigates some of its properties. Pre-topological rough sets is one of the topological generalizations of topological rough sets and it is the details of Section 5. Section 6 studied the relative topological rough classes. The conclusion work appears in Section 7.

2. Basics of Pawlak rough sets

The following illustrations about rough sets can be found in [18-23, 34-41].
Definition 2.1 Let $A = (U, R)$ be an approximation space. The equivalence classes U/R of the relation equivalence relation R will be called elementary sets (atoms) in A.
Definition 2.2 Let $A = (U, R)$ be an approximation space. Every finite union of elementary sets in A will be called a composed set in A. The family of all composed sets in A will be denoted by com (A). The family com (A) in the approximation space $A = (U, R)$ is a topology on the set U.

Definition 2.3 Let A = (U, R) be an approximation space; and let X be a certain subset of U. The least composed set in A containing X will be called the best upper approximation of X in A, and is denoted by $\overline{R}(X)$. The greatest composed set in A contained in X will be called the best lower approximation of X in A, and is denoted by $\underline{R}(X)$.

Definition 2.4 Let X is a subset of an approximation space A = (U, R). The set $Bnd(X)=\overline{R}(X)-\underline{R}(X)$ is called the boundary of X in A. The sets $\underline{Edg}(X)=X-\underline{R}(X)$ and $\overline{Edg}(X)=\overline{R}(X)-X$ are referred to as an internal and external edges of X in A, respectively.

For a subset X of an approximation space A = (U, R), we have $Bnd(X)=\overline{Edg}(X)\cup\underline{Edg}(X)$.

Since the approximation space A = (U, R), defines uniquely the topological space $\tau_R=(U,com(A))$, and com(A) is the family of all open sets in τ_R, and U/R is a basis for τ_R, then τ_R is a quasi-discrete topology on U, and com(A) is both the set of all open and closed sets in τ_R. Thus, $\underline{R}(X)$ and $\overline{R}(X)$ can be interpreted as the interior and the closure of the set X in the topological space τ_R, respectively.

Definition 2.5 Let A = (U, R) be an approximation space, and let X be a subset of U. The set X is called rough in A if and only if $\overline{R}(X)\neq\underline{R}(X)$, otherwise, X is an exact set in A.

Example 2.1 Let $U=\{a,b,c,d\}$ and $R=\{(a,a),(a,c),(c,a),(b,b),(c,c),(d,d)\}$ be an equivalence relation on U, then A = (U, R) is an approximation space. The set of atoms of A is $U\backslash R=\{\{a,c\},\{b\},\{d\}\}$, for $X=\{a,d\}$ we have $\underline{R}(X)=\{d\}$ and $\overline{R}(X)=\{a,c,d\}$, i.e., $\underline{R}(X)\neq\overline{R}(X)$ in A, then X is a rough set in A = (U, R).

Some properties of the approximations are given in the following proposition.

Proposition 2.1 Let A = (U, R) be an approximation space, and let X and Y be subsets of U, then we have:

(i) $\underline{R}(X)\subset X\subset\overline{R}(X)$.

(ii) $\underline{R}(U)=\overline{R}(U)=U$, $\underline{R}(\phi)=\overline{R}(\phi)=\phi$.

(iii) $\overline{R}(\overline{R}(X))=\underline{R}(\overline{R}(X))=\overline{R}(X)$.

(iv) $\underline{R}(\underline{R}(X))=\overline{R}(\underline{R}(X))=\underline{R}(X)$.

(v) $\overline{R}(X\cup Y)=\overline{R}(X)\cup\overline{R}(Y)$.

(vi) $\underline{R}(X\cap Y)=\underline{R}(X)\cap\underline{R}(Y)$.

(vii) $\overline{R}(X\cap Y)\subset\overline{R}(X)\cap\overline{R}(Y)$.

(viii) $\underline{R}(X\cup Y)\supset\underline{R}(X)\cup\underline{R}(Y)$.

(ix) $\overline{R}(X-Y)\supset\overline{R}(X)-\overline{R}(Y)$

(x) $\underline{R}(X-Y)\subset\underline{R}(X)-\underline{R}(Y)$

3. Topological rough sets and its properties

The main purpose of this section is to point out that the concept of rough sets have a purely topological nature [26-29,42]. At the same time a more general notion of a topological rough set will be considered, which has an independent interest, since its construction depends on a different approach than that due to Pawlak in [18].

Definition 3.1 Let (U,τ) be a topological space. The topology τ defines the equivalence relation R_τ on the power set P (U) given by the condition:

$(X,Y)\in R_\tau$ Iff $\text{int}(X)=\text{int}(Y)$ and $cl(X)=cl(Y)$.

Definition 3.2 Let (U,τ) be a topological space, and X, Y be subsets of U. The class $RO(U)$ of subsets of U is called a topological rough class if for every X and Y in $RO(U)$, we have $\text{int}(X)=\text{int}(Y)$ and $cl(X)=cl(Y)$.

Example 3.1 Let $U=\{a,b,c\}$ with the topology $\tau=\{U,\phi,\{a\},\{a,b\}\}$, then the equivalence classes of the equivalence relation R_τ are $P(U)/R_\tau=\{\{\{a\},\{a,c\}\},\{\{b\},\{b,c\}\},\{\{c\},\{a,b\}\}\}$.

Then each element in $P(U)/R_\tau$ is a topological rough set.

Example 3.2 Let R be an equivalence relation on a set U; let τ be the set of all subsets X of U, such that if $x\in X$ and $(x,y)\in R$ then $y\in X$.

then τ is a quasi-discrete topology on U, and topological rough sets in the topological space (U,τ) are the same as Pawlak rough sets in the approximation space A = (U, R).

Example 3.3 Let $U=\{a,b,c,d,e\}$ with the equivalence relation $R=\{(a,a),(b,b),(a,b),(b,a),(c,c),(d,d),(e,e)\}$. Then the equivalence classes of R are $U\backslash R=\{\{a,b\},\{c\},\{d\},\{e\}\}$. Then the quasi-discrete topology generated by the equivalence relation R is $\tau_R = \{U,\phi,\{c\},\{d\},\{e\},\{a,b\},\{c,d\},\{c,e\},\{d,e\},\{a,b,c\},\{a,b,d\},\{a,b,e\},\{c,d,e\},\{a,b,d,e\},\{a,b,c,e\},\{a,b,c,d\}\}$.

Extending Pawlak's terminology and according to Example 3.3, we have:

(i)Any subset **X** of **U** and $X \notin \tau_R$ is a rough set, for instance, the subset $X = \{a,c\}$ such that $\text{int}(X) = \{c\}$ and $cl(X) = \{a,b,c\}$.

(ii)The element **c** in the rough set $X = \{a,c\}$ surely belongs to {a, c} since $c \in \text{int}\{a,c\}$, but the element **a** possibly belongs to {a, c} for $a \in cl\{a,c\}$.

Lemma 3.1 For any topology τ on a set **U**, and for all x, y in **U**, the condition $x \in cl(\{y\})$ and $y \in cl(\{x\})$ implies $cl(\{x\}) = cl(\{y\})$.

Lemma 3.2 If τ is a quasi-discrete topology on a set **U**, then $y \in cl(\{x\})$ implies $x \in cl(\{y\})$ for all $x, y \in U$.

Proposition 3.1 If τ is a quasi-discrete topology on a set **U**, then the family $\{cl(\{x\}) : x \in U\}$ is a partition of **U**.

Proof: Suppose that $C = \{cl(\{x\}) : x \in U\}$, then

(i) Since $x \in cl(\{x\})$ for all $x \in U$, then $\bigcup_{x \in U} cl(\{x\}) = U$.

(ii) For any $x, y \in U$, either $cl(\{x\}) = cl(\{y\})$ or $cl(\{x\}) \bigcap cl(\{y\}) = \phi$, such that if $cl(\{x\}) \bigcap cl(\{y\}) \neq \phi$ then there is an element $z \in U$ such that $z \in cl(\{x\})$ and $z \in cl(\{y\})$, then by Lemma 3.1 and Lemma 3.2, we have $x \in cl(\{z\})$ and $y \in cl(\{z\})$, and we have $cl(\{z\}) = cl(\{x\}) = cl(\{y\})$, hence $cl(\{x\}) = cl(\{y\})$ or $cl(\{x\}) \bigcap cl(\{y\}) = \phi$ for every $x \neq y$, hence **C** is a partition of the set **U**.

Proposition 3.2 For every quasi-discrete topology τ on a set **U**, there is an equivalence relation **R** on **U**, such that a subset **X** of **U** is open in (U,τ) if and only if $x \in X$, $(x,y) \in R$, then $y \in X$.

Proof: The relation **R** defined by $R = \{(x,y) : x,y \in U$ and $x \in cl(\{y\})\}$ is an equivalence relation satisfies the condition of the proposition.

Theorem 3.1 Pawlak rough sets are exactly the same as rough sets in a quasi-discrete topology.

Proof: Follows from Proposition 3.2.

In the following we give an example to illustrate the above theorem.

Example 3.4 Let $U = \{a,b,c,d\}$ with the equivalence relation $R = \{(a,a),(b,b),(a,c),(c,a),(c,c),(d,d)\}$. Then $\mathsf{A} = (U,R)$ is an approximation space, and $U / R = \{\{a,c\},\{b\},\{d\}\}$ is the class of elementary sets of **R** in **A**. And for $X \subset U$, such that $X = \{b,c\}$, we have $\underline{R}(X) = \{b\}$ and $\overline{R}(X) = \{a,b,c\}$, i.e. the subset **X** is a Pawlak rough set in the approximation space $\mathsf{A} = (U,R)$. Now the topology on **U** generated by **R**, which has U/R as a base is given by: $\tau_R = \{U,\varnothing,\{b\},\{d\},\{a,c\},\{b,d\},\{a,b,c\},\{a,c,d\}\}$. Then for the same subset $X = \{b,c\}$, we have $\text{int}(X) = \{b\}$ and, $cl(X) = \{a,b,c\}$ i.e. a rough set in the topological space (U,τ_R). Hence Pawlak rough sets are the same as rough sets for a quasi-discrete topology τ_R.

In the classical papers of Pawlak [18-23] , he mentioned the term rough set in two places with different meaning.

1-The subset **X** of the approximation space $\mathsf{A} = (U,R)$ is called a rough set if $\underline{R}(X) \neq \overline{R}(X)$.

2-If $\mathsf{A} = (U,R)$ is an approximation space, then the relation $\approx$ defined by $X \approx Y$ iff $\underline{R}(X) = \underline{R}(Y)$ and $\overline{R}(X) = \overline{R}(Y)$ is an equivalence relation on the power set P(U). An equivalence class of $P(U) / \approx$ is called a rough set [18].

Wiweger in [29] followed Pawlak and used the same terminology. We see that it is useful to avoid confusion between the two concepts by replacing the term topological rough set by topological rough class and the rough set defined in [18] by Pawlak rough class.
Accordingly, Pawlak rough classes are exactly the same as topological rough classes and Pawlak rough sets are exactly the same as rough sets for a quasi-discrete topology.

4. An alternative description of topological rough classes

Definition 4.1 Let (U,τ) be a topological space. The rough pair in (U,τ) is a pair (M,N), where **M** and **N** are subsets of **U**, satisfying the following four conditions

(1) M is an open set i.e., $M \in \tau$.

(2) N is a closed set i.e., $N \in \tau^C$.

(3) $M \subset N$.

(4) The set $N - cl(M)$ contains a subset Z that satisfies the conditions $\text{int}(Z) = \phi$ and $N - cl(M) \subset cl(Z)$.

Proposition 4.1 For any subset X of a topological space (U,τ), the pair $(\text{int}(X), cl(X))$ is a rough pair in (U,τ).

Proof : Let $M = \text{int}(X)$ and $N = cl(X)$, then $M \in \tau$ and $N \in \tau^C$. $M \subset N$, Such that $\text{int}(X) \subset X \subset cl(X)$. Define the subset $Z = X - cl(M)$, Since $X \subset cl(X) = N$, then $X - cl(M) \subset N - cl(M)$ i.e. $Z \subset N - cl(M)$. For $\text{int}(Z) = \phi$, suppose that $\text{int}(Z) \neq \phi$, then there exist an open set G such that $G \subset Z = X - cl(M)$ i.e. $G \subset X$ and $G \not\subset cl(M)$, i.e. $G \subset X$ and $G \not\subset M$, i.e. $G \subset X$ and $G \not\subset \text{int}(X)$, and this contradiction, then must $\text{int}(Z) = \phi$. Finally, for $N - cl(M) \subset cl(Z)$ let $a \in N - cl(M)$ then $a \in N$ and $a \notin cl(M)$. If $a \in X$, then $a \in Z$. Hence $a \in cl(Z)$. If $a \notin X$ and $a \in N = cl(X)$ and $a \notin cl(M)$, let G be an open set containing a, i.e. $a \in G$, but $a \in cl(X)$, then by definition of $cl(X), G \bigcap X \neq \phi$, but $a \in G - cl(M) = G \bigcap [cl(M)]^c$ which is an open set containing a, then $G \bigcap [cl(M)]^c \bigcap X \neq \phi$ hence $G \bigcap [\ X - cl(M)\ ] \neq \phi$. Hence $G \bigcap Z \neq \phi$ and $a \in cl(Z)$. Consequently $N - cl(M) \subset cl(Z)$.

Proposition 4.2 For any rough pair (M,N) in the topological space (U,τ), there is a subset X of U, such that $M = \text{int}(X)$ and $N = cl(X)$.

Proof : Let (M,N) be a rough pair and let Z be a subset of U, satisfies that $\text{int}(Z) = \phi$, $Z \subset N - cl(M)$ and $N - cl(M) \subset cl(Z)$. Define $X = M \bigcup Z$, then:

For $M = \text{int}(X)$, we prove that $M \subset \text{int}(X)$ and $\text{int}(X) \subset M$. For $M \subset \text{int}(X)$, since $X = M \bigcup Z$, then $M \subset X$, hence $\text{int}(M) \subset \text{int}(X)$ but $M = \text{int}(M)$, hence $M \subset \text{int}(X)$. For $\text{int}(X) \subset M$ let $a \in \text{int}(X)$, hence $a \in X$ i.e. $a \in M \bigcup Z$ and we have three cases:

(I) $a \in M$ or (II) $a \in Z$ or (III) $a \in M \bigcap Z$: We prove now that these three cases are two only by proving that $M \bigcap Z = \phi$. To this end suppose that $M \bigcap Z \neq \phi$. Then there is an element x such that $x \in M$ and $x \in Z$. But $Z \subset N - cl(M)$. Then $x \in N$ and $x \notin cl(M)$ i.e. $x \notin M$, i.e. $x \in M$ and $x \notin M$, and this contradiction. Hence $M \bigcap Z = \phi$. Now $X = M \bigcup Z$, then $M = X - Z$ and $Z = X - M$. Now if (I) $a \in M$ holds then $\text{int}(X) \subset M$, and if (II) $a \in Z$ holds, then $a \in X - M$ i.e. $a \in X$ and $a \notin M$, hence $a \notin X - Z$ i.e., $a \notin X$, hence $a \notin \text{int}(X)$, i.e. if $a \notin M$ we have $a \notin \text{int}(X)$ and this proves that $\text{int}(X) \subset M$.

(2) For $N = cl(X)$. Since $X = M \bigcup Z$, then $cl(X) = cl(M \bigcup Z)$ i.e. $cl(X) = cl(M) \bigcup cl(Z)$, but $M \subset N$ i.e. $cl(M) \subset cl(N)$, hence $cl(X) \subset cl(N) \bigcup cl(Z)$ i.e. $cl(X) \subset cl(N \bigcup Z)$, but $Z \subset N$ such that $Z \subset N - cl(M)$, then $cl(X) \subset cl(N) = N$.

On the other hand,

$$N = cl(N) = cl\ [\, cl(M) \bigcup (N - cl(M)\,]$$
$$= cl(cl(M)) \bigcup cl(N - cl(M))$$
$$= cl(M) \bigcup cl(N - cl(M)).$$

But $N - cl(M) \subset cl(Z)$, then $cl(N - cl(M)) \subset cl(Z)$. Consequently $N \subset cl(M) \bigcup cl(Z) = cl(M \bigcup Z) = cl(X)$.

Example 4.1 Let $U = \{a,b,c,d,e\}$ with the topology $\tau = \{U, \phi, \{a\}, \{c\}, \{e\}, \{c,e\}, \{b,e\}, \{a,c\}, \{a,e\}, \{a,c,e\}, \{b,c,e\}, \{a,b,e\}, \{a,b,c,e\}\}$.

Then for the subset $X = \{d,e\}$ we have $\text{int}(X) = \{e\}$ and $cl(X) = \{b,d,e\}$, then $(\text{int}(X), cl(X)) = (\{e\}, \{b,d,e\})$ is a rough pair in (U,τ).

Example 4.2 Let $U = \{a,b,c,d\}$ with the topology $\tau = \{U, \phi, \{a\}, \{c\}, \{a,c\}, \{b,d\}, \{a,b,d\}, \{c,b,d\}\}$. Then the pairs $(\{a\}, \{a,b,d\})$ and $(\{c\}, \{c,b,d\})$ are rough pairs in (U,τ), but the pair $(\{c\}, \{a,c\})$ is not a rough pair in (U,τ).

Theorem 4.1 For any topological space (U,τ), the pair (M, N) is a rough pair iff there exists a subset X of U such that M = int(X) and N = cl (X).

Proof: The proof is given from Propositions 4.1 and 4.2.

The subset X, corresponding to the rough pair $(\text{int}(X),cl(X))$ is not unique in the topological space (U,τ), as illustrated in the following example.

Example 4.3 Let $U=\{a,b,c,d\}$ with the topology $\tau=\{U,\tau,\{a\},\{c\},\{a,c\},\{b,d\},\{a,b,d\},\{b,c,d\}\}$. For the rough pair $(\{a\},\{a,b,d\})$ there are two subsets $X=\{a,b\}$ and $Y=\{a,d\}$ such that $\text{int}(X)=\text{int}(Y)$ and $cl(X)=cl(Y)$.

Definition 4.2 Let (U,τ) be a topological space; and let the class $RO(U)$ be a topological rough class in (U,τ). We define the function $f: P(U)/R_\tau \longrightarrow PA(U)$, by $f(X,Y)=(\text{int}(X),cl(X))$, and $X,Y\in RO(U)$, where $P(U)/R_\tau$ is the set of all topological rough classes in (U,τ) and $PA(U)$ the set of all rough pairs in (U,τ).

It follows from the above definition that $f(X,Y)$ does not depend on the choice of X and Y in the equivalence class $RO(U)$ as illustrated by the following example.

Example 4.4 Let $U=\{a,b,c\}$ with the topology $\tau=\{U,\phi,\{a\},\{a,b\}\}$ be a topological space. Then $P(U)/R_\tau=\{\{\{a\},\{a,c\}\},\{\{b\},\{b,c\}\},\{c\},\{a,b\}\}$.

Hence $f(\{\{a\},\{a,c\}\})=(\text{int}\{a\},cl\{a\})=(\text{int}\{a,c\},cl\{a,c\})=(\{a\},U)$.

Theorem 4.2 For any topological space (U,τ) the function $f((X,Y))=(\text{int}(X),cl(X))$ and $X\in RO(U)$ is a one to one and onto from the set of all topological rough classes onto the set of all rough pairs in (U,τ).

Proof: For the one to one part, let $(X1,Y1)$ and $(X2,Y2)$ be two topological rough classes in (U,τ), and let $f((X1,Y1))=f((X2,Y2))$. Then $(\text{int}(X_1),cl(Y_1))=(\text{int}(X_2),cl(Y_2))$. Hence $\text{int}(X_1)=\text{int}(X_2)$ and $cl(Y_1)=cl(Y_2)$ then $(X_1,X_2)\in R_\tau$, hence X_1,X_2 belong to the same element of $P(U)\setminus R_\tau$, hence $(X1,Y1)=(X2,Y2)$. For the onto part, Proposition 4.2 ends the proof.

5. Pre-topological rough classes

In Pawlak approximation space $A=(U,R)$, the topological space $(U,com(A))$ is generally a quasi-discrete space, in which $\text{int}\,cl(A)=cl(A)$ for any $A\subset U$. Thus every subset A in Pawlak space is pre-open. Using another space (U,τ), which is not quasi-discrete, so it is possible to use pre-open concepts since $PO(U,\tau)\neq P(U)$. In this space, for any $A\subset U$, we have $\text{int}(A)\subset p.\text{int}(A)\subset A\subset p.cl(A)\subset cl(A)$. This implies decomposition for the boundary region $(cl(A)-\text{int}(A))$, which enlarges the range of a membership and consequently helps in applications.

Definition 5.1 Let (U,τ) be a topological space. Then we define the equivalence relation η_τ on the set P(U), by $(X,Y)\in\eta_\tau$ if $p\cdot\text{int}(X)=p\cdot\text{int}(Y)$ and $p\cdot cl(X)=p\cdot cl(Y)$.

Definition 5.2 Let (U,τ) be a topological space, the class Γ is called a pre-topological rough class if it contains all subsets X and Y of U, with $p\cdot\text{int}(X)=p\cdot\text{int}(Y)$ and $p\cdot cl(X)=p\cdot cl(Y)$.

Example 5.1 Let $U=\{a,b,c,d\}$, with topology

$$\tau=\begin{Bmatrix}U,\phi,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\\ \{b,c\},\{a,b,c\}\end{Bmatrix}$$

Then the pre-topological rough classes in (U,τ) are

$$P(U)/\eta_\tau=\{\{\{a\},\{a,d\}\},\{\{b\},\{b,d\}\},\{\{c\},\{c,d\}\},\{\{\phi\},\{d\}\},\{\{a,b\},\{a,b,d\}\},\{\{a,c\},\{a,c,d\}\},\{\{b,c\},\{b,c,d\}\},\{\{a,b,c\},\{U\}\}$$

The set $P(U)/\eta_\tau$ of all pre-topological rough classes in a topological space (U,τ) is a partition of the set $P(U)$.

6. Relative topological rough classes

Definition 6.1 Let (U,τ) be a topological space, and for $X \subset U$ let (X,τ_X) be a subspace of (U,τ). The topology τ_X on X defines the equivalence relation R_{τ_X} on the power set P (X) such that $(X1, X2) \in R_{\tau_X}$ iff $\text{int}_{\tau_X}(X1) = \text{int}_{\tau_X}(X2)$ and $cl_\tau(X_1) = cl_\tau(X_2),\ for\ X_1, X_2 \in P(X)$.

Definition 6.2 Let (U,τ) be a topological space, and let (X,τ_X) be a subspace of (U,τ), where $X \subseteq U$. The collection $P(X)/R_{\tau_X}$ is a partition of P (X), and any class $\Gamma \in P(X)/R_{\tau_X}$ is called a relative topological rough class.

Definition 6.3 Let (U,τ) be a quasi-discrete topological space, and let (X,τ_X) be a subspace of (U,τ), where $X \subseteq U$ is an open set in (U,τ). If $P(U)/R_\tau$ is the collection of all topological rough classes in (U,τ), then we call the collection $[P(U)/R_{\tau_X}]_X = \{[X \cap W] : [W] \in P(U)/R_\tau\}$ the sub-topological rough classes in (U,τ).

Proposition 6.1 Let (U,τ) be a quasi-discrete space, (X,τ_X) be a subspace of (U,τ), and let X be open in (U,τ). Then $P(X)/R_{\tau_X} = [P(U)/R_\tau]_X$.

Proof: We can rewrite the above equality as follows:

$$[P(U)/R_{\tau_X}]_X = \{[X \cap W] : [W] \in P(U)/R_\tau\}.$$

Now, let $[A] \in P(X)/R_{\tau_X}$, then there are subsets $A_1,\ldots,A_n$ of X such that $\text{int}_{\tau X}(A_1) = \text{int}_{\tau X}(A_2) = \cdots = \text{int}_{X\tau}(A_n)$, and $cl_{\tau X}(A_1) = cl_{\tau X}(A_2) = \cdots = cl_{\tau X}(A_n)$ but each τ_X-open is also τ_X-closed, such that (U,τ) is a quasi-discrete space, since X is open and closed in (U,τ), then

$\text{int}(A_1) \cap X = \text{int}(A_2) \cap X = \cdots = \text{int}(A_n) \cap X$ and

$cl(A_1) \cap X = cl(A_2) \cap X = \cdots = cl(A_n) \cap X$ i.e., there is $[W] \in P(U)/R_\tau$, such that $X \cap [W] \in [P(U)/R_\tau]_X$ and $[A] = [W] \cap X$.

The following example study the cases when the topology in the above proposition is not a quasi-discrete topology, and the subset X of U is open and study the case when X is not open.

Example 6.1 Let (U,τ) be a topological space, where $U = \{a,b,c,d\}$ and $\tau = \{U, \phi, \{c\}, \{a,b\}, \{a,b,c\}\}$, let $X \subset U$ such that $X = \{b,c,d\}$, then $\tau_X = \{X, \phi, \{c\}, \{b\}, \{b,c\}\}$ and (X,τ_X) is the subspace of (U,τ), then the relative topological rough classes in (U,τ) are:

$$P(X)/R_{\tau_X} = \{\{\{b\},\{b,d\}\},\{\{c\},\{c,d\}\}, \{\{d\}\},\{\{b,c\}\}\},$$

But the topological rough classes are:

$$P(U)/R_\tau = \{\{\{a\},\{b\},\{a,d\},\{b,d\}\},\{\{c\},\{c,d\}\},\{\{d\}\},\{\{a,b\},\{a,b,d\}\} \{\{a,b,c\}\},\{\{a,c\},\{b,c\},\{b,c,d\}\},\{\{a,c,d\}\}\}.$$

Then the sub-topological rough classes in (U,τ) are:

$$[P(U)/R_\tau]_X = \{\{\{b\},\{d\},\{b,d\}\},\{\{c\},\{c,d\}\},\{\{d\}\},\{\{b\},\{b,d\}\} \{\{b,c\}\},\{\{c\},\{b,c\},\{b,c,d\}\},\{\{c,d\}\}\}$$

We observe that $P(X)/R_{\tau X} \subset [P(U)/R_\tau]_X$ where τ is not a quasi-discrete topology and $X \notin \tau$. Also, we observe that $[P(U)/R_\tau]_X$ is not a partition of X. If $Y \in \tau$, say $Y = \{a,b,c\}$, then (Y,τ_Y) is a subspace of (U,τ), and in this case the relative topological rough classes are :

$$P(Y)/R_{\tau Y} = \{\{\{a\},\{b\}\},\{\{c\}\},\{\{a,b\}\}, \{\{a,c\},\{b,c\}\}\}.$$

The sub-topological rough classes are:

$$[P(U)/R_\tau]_Y = \{\{\{a\},\{b\}\},\{\{c\}\}, \{\{a,b\}\},\{\{a,c\},\{b,c\}\}\}.$$

Then we observe that: $P(Y)/R_{\tau_Y} \subset [P(U)/R_\tau]_Y$.

The following example study the case when the topology τ is a quasi-discrete and the subset X is open and is not open.

Example 6.2 Let (U,τ) be a quasi-discrete space, where $U = \{a,b,c,d\}$ and

$\tau = \{U, \phi, \{a\}, \{d\}, \{a,d\}, \{b,c\}, \{b,c,d\}, \{a,b,c\}\}$, then the topological rough classes in (U, τ) are

$$P(U)/R_\tau = \{\{\{a\}\}, \{\{b\},\{c\}\}, \{\{d\}\}, \{\{a,b\},\{a,c\}\}, \{\{a,d\}\}, \{\{b,c\}\}, \{\{b,d\}, \{c,d\}\}, \{\{a,b,c\}\}, \{\{b,c,d\}\}, \{\{a,b,d\}, \{a,c,d\}\}\}.$$

Let $X \subset U$ such that $X = \{a,b,c\}$, then $\tau_X = \{X, \phi, \{a\}, \{b,c\}\}$ and the sub-topological rough classes are:

$$[P(U)/R_\tau]_X = \{\{\{a\}\}, \{\{b,c\}\}, \{\{b\},\{c\}\}, \{\{a,b\},\{a,c\}\}\},$$

but the relative topological rough classes are

$$P(X)/R_{\tau_X} = \{\{\{a\}\}, \{\{b\},\{c\}\}, \{\{a,b\}, \{a,c\}\}, \{\{b,c\}\}\},$$

then we have $P(X)/R_{\tau_X} = [P(U)/R_\tau]_X$, for $X \in \tau$.

If $X^* \notin \tau$, say $X^* = \{a,c,d\}$, then (X^*, τ^*) is a subspace of (U, τ), such that $\tau^* = \{X^*, \phi, \{a\}, \{d\}, \{c\}, \{a,d\}, \{a,c\}, \{c,d\}\}$, then the sub-topological rough classes are:

$$[P(U)/R\tau]_{X^*} = \{\{\{a\}\}, \{\{c\}\}, \{\{d\}\}, \{\{a\}, \{a,c\}\}, \{\{a,d\}\}, \{\{d\}, \{c,d\}\}, \{\{a,c\}\}, \{\{c,d\}\}, \{\{a,d\}, \{a,c,d\}\}\}.$$

and the relative topological rough classes are:

$$P(X^*)/R_{\tau^*} = \{\{\{a\}\}, \{\{c\}\}, \{\{d\}\}, \{\{a,d\}\}, \{\{a,c\}\}, \{\{c,d\}\}\},$$

then we have $P(X^*)/R_{\tau^*} \subset [P(U)/R_\tau]_{X^*}$.

Let (U, τ) be a topological space. If (X, τ_X) is a subspace of (U, τ), then

If τ is not a quasi-discrete space, and $X \subseteq U$, then $P(X)/R_{\tau X} \subset [P(U)/R_\tau]_X$.

If τ is a quasi-discrete space, and $X \in \tau$, then $P(X)/R_{\tau X} = [P(U)/R_\tau]_X$.

But if τ is a quasi-discrete space, and $X \notin \tau$, then $P(X)/R_{\tau X} \subset [P(U)/R_\tau]_X$. The class $[P(U)/R_\tau]_X$ is not a partition of X in general.

7. Topological Rough sets with tools for data mining

The great advances in information technology have made it possible to store a great quantity of data. In the late nineties, the capabilities of both generating and collecting data were increased rapidly. Millions of databases have been used in business management, government administration, scientific and engineering data management, as well as many other applications. It can be noted that the number of such databases keeps growing rapidly because of the availability of powerful database systems. This explosive growth in data and databases has generated an urgent need for new techniques and tools that can intelligently and automatically transform the processed data into useful information and knowledge. One of the processes used to transform data into knowledge is knowledge discovery in database.

Rough sets and topological rough sets constitute a consistency base for data mining; they offers useful tools for discovering patterns hidden in data in many aspects. Although rough set approach deals with discrete data, rough set is commonly used in conjunction with other techniques connected to coded on the dataset.

Data mining technology provides a new thought for organizing and managing tremendous data. Topological rough sets are one of the important tools for knowledge discovery and rule extraction. This tool can used to analyze intact data, obtain uncertain knowledge and offer an effective tool by reasoning.

8. Conclusion

We conclude that the emergence of topology in the construction of some rough set concepts will help to get rich results that yields a lot of logical statements which discover hidden relations between data and moreover, probably help in producing accurate programs. Also, the topological rough generalizations of rough sets will help to define topological rough measures which can used to finding the attribute missing values in information systems.

References

[1] D. Andrijevic : On the topology generated by preopen sets, Mat. Vesnik 39 (1987), 367-376.

[2] M. Chuchro : On rough sets in topological Boolean algebras, RSKD '1993 Proceedings of the International Workshop on Rough Sets and Knowledge Discovery: Rough Sets, Fuzzy Sets and Knowledge Discovery; 157-160.

[3] P. C. W. Davies , J. Brown : Superstring, a theory of everything, Cambridge University Press. (2000).

[4] M. Eggen , D. Smith , S. A. Richard: A transition to advanced Mathematics, California (1990).

[5] A. S. Farrag , A. A. Sewisy: Computer construction and enumeration of topology on finite sets. Inter. J. computer Math. vol. 72, PP. 433-440, (1999).

[6] E. Flapan : When topology meets Chemistry, Cambridge University Press. (2000).
[7] M. Gehrke ; E. Walker : On the structure of rough sets, Bull. Polish Acad. Sci., Math., Vol. 40, No. 3, 1992, 235-245.
[8] T. Iwinski : Algebraic approach to rough sets. Bull. Polish Acad. Sci., Math., 35, pp. 673-683 (1987).
[9] T. Y. Kong , R. Kopperman and P. R. Meyer: A topological approach to digital topology. American Mathematical Monthly 98, , (1991) 901-917.
[10] E. Konrad , E. Orlowska, Z. Pawlak: An approximate concept learning, Berlin, Bericht, 1981, pp. 81-87.
[11] J. Kortelainen : On relationship between modified sets, topological spaces and rough sets, Fuzzy sets and systems, 61(1994) 91-95.
[12] I. Kramosil : An alternative approach to rough sets, Kyberrietika, Vol. 28, (1992), No. 1, 1-25.
[13] P. Maritz : Pawlak and topological rough sets in terms of multifunctions, Glasnik *Mathematiĉki* , Vol. 31(51), (1996), 77-99.
[14] A. S. Mashhour: New types of topological spaces, Pub. Math. Debrecen, 17 (1970), 77-80.
[15] J. R. Munkres : Topology a first course, Prentice-Hall, New Jersey (1975).
[16] A. Obtulowicz : Rough sets and Heyting algebra valued sets. Bull. Polish Acad. Sci., Math., 35, pp. 667-673 (1988).
[17] E. Orlowska : Algebraic approach to data base constrains. Fundamenta Informaticae, 10, pp. 57-68 (1987).
[18] Z. Pawlak, W. Marek : Rough sets and information systems, ICS. PAS. Reports (441) (1981), 481-485.
[19] Z. Pawlak : Rough sets, ICS. PAS. Reports (431) (1981), 311-341.
[20] Z. Pawlak : Rough sets, algebraic and topological approach, ICS. PAS. Reports (482) (1982), 99-104.
[21] Z. Pawlak : Rough sets, Int. Journal of computer and information sciences, vol. 11. No. 5 (1982), 341-356.
[22] Z. Pawlak : On rough relations, Bull. Pol. Ac. Tec. sciences, vol. 34. 9-10 (1986).
[23] Z. Pawlak : Rough sets, rough relations and rough functions Bull. Pol. Ac. Math, 15-19 (1996).
[24] A. Skowron : The relationship between rough set theory and evidence theory. Bull. Polish. Acad. Sci., Math., 37, (1989) 87-90.
[25] A. S. Salama , H. M. Abu-Donia ; New Approaches for Data Reduction, international journal of Pure and Applied Mathematics , Vol. 27,No. 1 , pp. 39-53(2006).
[26] A. S. Salama ; Topologies Induced by Relations with Applications, Journal of Computer Science 4(10): 879-889, (2008).
[27] A. S. Salama ; Two New Topological Rough Operators, Journal of Interdisciplinary Mathematics Vol. 11, No.1, New Delhi Taru Publications-, INDIA pp. 1-10(2008).
[28] A. S. Salama ; Topological Solution for missing attribute values in incomplete information tables, Information Sciences, 180 (2010) 631-639 .
[29]A. Wiweger, : On topological rough sets, Bull. Pol. Ac. Mat., 37 (1989) , 51-62.
[30] U. Wybraniec-Skardowski : On a generalization of approximation space. Bull. Pol. Acad. Sci., Math., 37, (1989) 51-62.
[31] W.Z. Wu, W.X. Zhang, Constructive and axiomatic approaches of fuzzy approximation operators, Information Sciences 159 (2004) 233–254.
[32] W.Z. Wu, Y. Leung, J.S. Mi, On characterizations of (I, T)-fuzzy rough approximation operators, Fuzzy Sets and Systems 154 (2005) 76–102.
[33] W.Z. Wu, W.X. Zhang, Rough set approximations vs. measurable spaces, in: IEEE International Conference on Granular Computing, 2006, pp. 329–332.
[34] W.H. Xu, W.X. Zhang, Measuring roughness of generalized rough sets induced by a covering, Fuzzy Sets and Systems 158 (2007) 2443–2455.
[35] J.T. Yao, A ten-year review of granular computing, in: Proceedings of 2007 IEEE International Conference on Granular Computing, 2007, pp. 734–739.
[36] Y.Y. Yao, Two views of the theory of rough sets in finite universes, International Journal of Approximate Reasoning 15 (1996) 291–317.
[37] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences 111 (1–4) (1998) 239–259.
[38] Y.Y. Yao, Constructive and algebraic methods of theory of rough sets, Information Sciences 109 (1998) 21–47.
[39] Y.Y. Yao, T.Y. Lin, Generalization of rough sets using modal logic, Intelligent Automation and Soft Computing, An International Journal 2 (1996) 103–120.
[40] H. Zhang, H. Liang, D. Liu, Two new operators in rough set theory with applications to fuzzy sets, Information Sciences 166 (1–4) (2004) 147–165.
[41] W. Zhu, F.Y. Wang, Reduction and axiomization of covering generalized rough sets, Information Sciences 152 (2003) 217–230.
[42] W. Zhu, Topological approaches to covering rough sets, Information Sciences 177 (6) (2007) 1499–1508.
[43] W. Zhu, F.Y. Wang, On three types of covering rough sets, IEEE Transactions on Knowledge and Data Engineering 19 (8) (2007) 1131–1144.
[44] W. Zhu, Generalized rough sets based on relations, Information Sciences 177 (22) (2007) 4997–5011.
[45] W. Zhu, Relationship between generalized rough sets based on binary relation and covering, Information Sciences 179 (2009) 210–225.

A . S. Salama received the BS degree in mathematics from Tanta University, Tanta, Egypt, in 1998 and the MS degree in Topological Rough Sets from the University of Tanta, Egypt, in 2002. He worked at Tanta university from 1998 to 2008. He studied as a PhD student at Tanta University, Tanta, Egypt, from 2002 to 2004, in Topology and Information Systems. He was assistant professor in the Department of Mathematics at College of Science in Dawadmi in King Saud University, KSA from 2008 until 2010. Currently, He is assistant professor at Shaqra University, KSA. His research interests include artificial intelligence, rough set, Data mining, Topology, Fuzzy sets, Information Systems.