

## Topologies Induced by Relations with Applications

A.S. Salama

Department of Mathematics, Faculty of Science, Tanta University, Egypt

---

**Abstract:** Topological structures induced by relations represents the process of extracting interesting decision rules from data. The attributes reduction and the calculation of core are the essentials to extract the decision rules. Finding the reducts, core and decision rules topologically is a new mathematical tool for discovering attribute dependencies in information systems. **Problem statement:** These mathematical tools employ some concepts from topological spaces, relational databases, rough sets and information systems. The results using our approach are more accurate and applicable than that using the classical approaches such as in the father's approach of rough sets (Pawlak approach). **Approach:** Topologies generated using dominance (pre-order) relations and general binary relation are the knowledgebase of our approximations. We suggested a new algorithm (openness algorithm) based on the topologies induced by general relations. **Results:** Results obtained by the proposed approach to find the reducts and core in terms of open and closed sets are compared with the existing method. Our proposed method is proved to be the accurate than results of any approaches using some types of binary relations such as order(pre-order) relations or symmetric relations. **Conclusion:** In this study, There are many approaches for obtaining topologies by relations and we used some of them in data reduction. These approaches were generalizations to Pawlak approaches namely, we ignored the notion of equivalence relations. Also, these approaches open the way for other approximations if we use the general topological recent concepts such as pre-open sets or semi-open sets.

**Key words:** Topological spaces, binary relations, relational databases, information system, rough sets, data reduction

---

### INTRODUCTION

Topology is an important and interesting area of mathematics, the study of which will not only introduce you to new concepts and theorems but also put into context old ones like continuous functions. It is so fundamental that its influence is evident in almost every other branch of mathematics. This makes the study of topology relevant to all who aspire to be mathematicians whether their first love is algebra, analysis, category theory, chaos, continuum mechanics, dynamics, geometry, industrial mathematics, mathematical biology, mathematical economics, mathematical finance, mathematical modeling, mathematical physics, mathematics of communication, number theory, numerical mathematics, operation research or statistics. Topological notions like compactness, connectedness and denseness are as basic to mathematicians of today as sets and functions were to those of last century<sup>[3,8,9]</sup>.

For a long time, many individuals believed that abstract topological structures have limited application in the generalization of real line and complex plane or some connections to Algebra and other branches of

mathematics. And it seems that there is a big gap between these structures and real life applications. We noticed that in some situations, the concept of relation is used to get topologies that are used in important applications such as computing topologies<sup>[15]</sup>, recombination spaces<sup>[2,7,16]</sup> and information granulation which are used in biological sciences and some other fields of applications.

The aim of rough set theory is to give a description of the set of objects by logical, set-theoretical, topological etc. tools in terms of similarity relations and derived notions related by these relations. The description of the set of objects entails as well relationships and functional or near to functional dependencies among various similarity relations generated by various sets of the set of objects.

### MATERIALS AND METHODS

Rough sets were first introduced by Pawlak<sup>[10,11]</sup> and are based on approximation spaces. An approximation space is a pair  $A = (Ob, R)$ . Here,  $R$  is an equivalence relation, also called indiscernibility relation, imposing a granularity on the universe  $Ob$  such that  $R \subseteq Ob \times Ob$ .

Furthermore, we assume  $Ob$  to be finite. For  $x \in Ob$ , let  $[x]_R$  be the equivalence class containing  $x$ , i.e.,  $[x]_R = \{y: y R x\}$ .

Given an arbitrary set  $X \subseteq Ob$ , we wish to describe  $X$  in terms of elements or granules of  $Ob/R$ . Pawlak proposed the use of lower and upper approximations of a set  $X$ , denoted  $\underline{R}(X)$  and  $\overline{R}(X)$ , respectively. Lower and upper approximations are defined as:

$$\begin{aligned}\underline{R}(X) &= \{x \in Ob: [x]_R \subseteq X\} \\ \overline{R}(X) &= \{x \in Ob: [x]_R \cap X \neq \emptyset\}\end{aligned}$$

The semantics of the approximations of sets may be defined as follows:

- Elements of the universe that belong to  $\underline{R}(X)$  are those elements that surely belong to the set  $X$
- Elements that belong to  $\overline{R}(X)$  possibly belong to the set  $X$
- Elements that belong to  $Ob/\overline{R}(X)$  are elements of the universe that surely do not belong to the set  $X$ . Hence, the uncertainty lies in  $\overline{R}(X)/\underline{R}(X)$  which is also called area of uncertainty. Elements of the area of uncertainty may, or may not, belong to  $X$

The approximation operators can also be considered using membership functions. It is possible to define a rough membership function as presented in<sup>[12]</sup>.

**Topologies induced by relations:** Let  $A = (Ob, R)$  be an approximation space. The equivalence classes  $Ob/R$  of the relation  $R$  will be called elementary sets (atoms) in  $A$ . 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  $A = (Ob, R)$  is a topology on the set  $Ob$ .

Since the approximation space  $A = (Ob, R)$ , defines uniquely the topological space  $\tau(A) = (Ob, com(A))$  and  $com(A)$  is the family of all open sets in  $\tau(A)$  and  $U/R$  is a basis for  $\tau(A)$ , then  $\tau(A)$  is a quasi-discrete topology on  $Ob$  and  $com(A)$  is both the set of all open and closed sets in  $\tau(A)$ . Thus, the lower approximation and the upper approximation of any subset  $X \subseteq Ob$  can be interpreted as the interior and the closure of the set  $X$  in the topological space  $\tau(A)$ , respectively.

**Lemma 1:** If  $\beta$  is a base for a topological space  $(Ob, \tau)$ , where  $\beta$  is a partition of  $Ob$ , then for every subset  $X \subseteq Ob$ :

- $int_\tau(X) = \bigcup \{B \in \beta: B \subseteq X\}$
- $cl_\tau(X) = \bigcup \{B \in \beta: B \cap X \neq \emptyset\}$

**Proof:** Only we prove (ii) because (i) is trivial. Let  $x \in cl_\tau(X)$  then for every open set  $G$  containing  $x$ ,  $X \cap G \neq \emptyset$ . But  $G = \bigcup_{B \in \beta} B$ , then there exists  $B_0 \subseteq G$  such that  $x \in B_0 \subseteq G$ . But  $B_0$  is an open set containing  $x$ , hence  $B_0 \cap X \neq \emptyset$  and  $x \in \bigcup \{B \in \beta: B \cap X \neq \emptyset\}$ .

Conversely, if  $x \in \bigcup \{B \in \beta: B \cap X \neq \emptyset\}$  and  $G$  is an open set containing  $x$  and  $\beta$  is a partition of  $Ob$ ,  $x \in U$ , then  $x$  belongs to only one element of  $\beta$  say  $x \in B_0$ . Then must  $B_0 \subseteq G$ , i.e.,  $x \in B_0 \subseteq G$  but  $B_0 \cap X \neq \emptyset$ , hence  $G \cap X \neq \emptyset$ . Then  $x \in cl_\tau(X)$ .

Let  $A_1 = (Ob, R_1)$  and  $A_2 = (Ob, R_2)$  be two approximation spaces. Then we say that the partition  $Ob/R$  depends on the partition  $Ob/R_2$  denoted  $Ob/R_1$  if and only if  $B = \bigcup_{S \in Ob/R_2} S, \forall B \in Ob/R_1$ .

**Proposition 2:** Let  $\tau_1$  and  $\tau_2$  be the topologies induced by the partitions  $Ob/R_1$  and  $Ob/R_2$  respectively. Then  $Ob/R_1 \leq Ob/R_2$  iff  $\tau_2 \subseteq \tau_1$ .

**Example:** Consider the partitions  $\beta_1 = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}\}$  and  $\beta_2 = \{\{x_1, x_2\}, \{x_3, x_4\}\}$  of the set  $Ob = \{x_1, x_2, x_3, x_4\}$ . Then  $\beta_1 \leq \beta_2$  and  $\tau_2 \subseteq \tau_1$  where  $\tau_1 = \{Ob, \emptyset, \{x_3\}, \{x_4\}, \{x_{1,2}\}, \{x_3, x_4\}, \{x_1, x_2, x_3\}, \{x_1, x_2, x_4\}\}$ ,  $\tau_2 = \{Ob, \emptyset, \{x_1, x_2\}, \{x_3, x_4\}\}$  are the topologies generated by  $\beta_1$  and  $\beta_2$  respectively.

For any topological space  $(Ob, \tau)$ , we define the equivalence relation  $E(\tau)$  on the set  $Ob$  by  $(x, y) \in E(\tau)$  iff  $cl_\tau(\{x\}) = cl_\tau(\{y\}), \forall x, y \in Ob$ . The set of all equivalence classes of  $E(\tau)$  is denoted by  $Ob/E(\tau)$ .

**Proposition 3:** Let  $A = (Ob, R)$  be an approximation space and let  $\tau_R$  be the topology generated by the base  $B_R = Ob/R$ . If  $(Ob, \tau)$  is the quasi-discrete topological space has  $Ob/E(\tau)$  as a base. Then  $\tau_R = \tau$  iff for all  $x \in B_R \in \beta_R$  there exists  $B \in Ob/E(\tau)$  such that  $x \in B$ .

**Lemma 4**<sup>[15]</sup>: For any topology  $\tau$  on a set  $Ob$  and for all  $x, y \in Ob$ , if  $y \in cl_\tau(\{x\})$  and  $x \in cl_\tau(\{y\})$  then  $cl_\tau(\{x\}) = cl_\tau(\{y\})$ .

**Lemma 5**<sup>[15]</sup>: If  $\tau$  is a quasi-discrete topology on a set  $Ob$ , then  $y \in cl_\tau(\{x\})$  implies  $x \in cl_\tau(\{y\})$  for all  $x, y \in Ob$ .

**Lemma 6**<sup>[15]</sup>: If  $\tau$  is a quasi-discrete topology on a set  $Ob$ , then the family  $\{cl_\tau(\{x\}): x \in Ob\}$  is a partition of  $Ob$ .

**Proposition 7:** Let  $\tau$  be the topology induced by the partition  $\beta_R = Ob/R$ . Then  $\beta_R = Ob/E(\tau)$ .

**Proof:**  $x \in B, B \in \beta_R$

- iff  $x \in cl_\tau(B) = \bigcup_{y \in B} cl_\tau(\{y\})$
- iff  $y_0 \in B$  and  $x \in cl_\tau(\{y_0\})$  iff  $cl_\tau(\{x\}) = cl_\tau(\{y_0\})$  (Lemma 2.2)
- iff  $(x, y_0) \in E(\tau)$
- iff  $A \in Ob/E(\tau)$  such that  $x \in A$
- iff  $\beta_R = Ob/E(\tau)$

For any  $n$  approximation spaces  $A_1 = (Ob, R_1), A_2 = (Ob, R_2), \dots, A_n = (Ob, R_n)$  we define the partition  $Ob/E(\tau_{ind}) = \bigcap_{i=1,2,\dots,n} Ob/E(\tau_i)$ .

**Theorem 8:**  $\tau_i \subseteq \tau_{ind}, i = 1, 2, \dots, n$  where  $\tau_i$  and  $\tau_{ind}$  are the topologies generated by the partitions  $Ob/(\tau_i)$  and  $Ob/E(\tau_{ind})$  respectively.

**Proof:** Since  $Ob/E(\tau_{ind}) \leq Ob/E(\tau_i)$  for all  $i = 1, 2, \dots, n$  then  $\tau_i \subseteq \tau_{ind}$ .

**Example:** Consider the topological space  $(Ob, \tau)$  where  $Ob = \{x_1, x_2, x_3, x_4\}$  and  $b = \{\{x_1\}, \{x_2, x_3\}, \{x_4\}\}$  is the base of  $\tau$ , then  $\tau$  is a quasi-discrete topology and:

$$cl_\tau(\{x_1\}) = \{x_1\}, cl_\tau(\{x_2\}) = \{x_2, x_3\}, cl_\tau(\{x_3\}) = \{x_2, x_3\}$$

$$cl_\tau(\{x_4\}) = \{x_4\}$$

$$\text{Then } Ob/E(\tau) = \{\{x_1\}, \{x_2, x_3\}, \{x_4\}\} = \beta.$$

**Example:** Consider the approximation spaces  $A_1 = (Ob, R_1), A_2 = (Ob, R_2)$  and  $A_3 = (Ob, R_3)$  where  $Ob = \{x_1, x_2, x_3, x_4\}$  and  $Ob/E(\tau_1) = \{\{x_1\}, \{x_2, x_3\}, \{x_4\}\}, Ob/E(\tau_2) = \{\{x_1, x_2\}, \{x_3, x_4\}\}$  and  $Ob/E(\tau_3) = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}\}$  are the bases of  $\tau_1, \tau_2$  and  $\tau_3$  respectively, then  $Ob/E(\tau_{ind}) = (Ob/E(\tau_1)) \cap (Ob/E(\tau_2)) \cap (Ob/E(\tau_3)) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}\}$  is the partition induced by  $E(\tau_{ind})$ . Then  $\tau_i \subset \tau_{ind}, i = 1, 2, 3$ .

**Topologies generated using similarity relations:** A similar relation  $R$  on  $Ob$  is any relation satisfies:

- For any  $x \in Ob, xRx$  (reflexive)
- For any  $x, y \in Ob$ , if  $xRy$  then  $yRx$  (Symmetric)

For  $x \in Ob$ , we define the similar class containing  $x$  by  $R(x) = \{y \in Ob: xRy\}$ .

The relation  $R$  on  $Ob$  defined by  $xRy$  iff  $d(x,y) < n$  where  $(Ob, d)$  is a metric space with a metric function  $d$  defined as:  $d(x,y) = |x - y|$  and  $n = \text{card}(Ob)$ , is a similar relation.

**Proposition 9** For any similar relation  $R$  defined on  $Ob$  we have:

- $x \in R(x)$
- $y \in R(x)$  iff  $x \in R(y)$
- $xRy$  iff  $x \in R(y)$  and  $y \in R(x)$

The class  $\beta = \{B(x): x \in X\}$  is called a symmetric covering of a set  $X$  if  $x \in B(y)$  iff  $y \in B(x)$ . Then the class  $\beta = \{R(x): x \in Ob\}$  is a symmetric covering of the set of objects  $Ob$ .

Let  $\beta$  is the symmetric covering of  $Ob$  by the similar relation  $R$ . Then we define a relation  $R_\beta$  induced by  $\beta$  by  $x R_\beta y$  iff there exist  $B \in \beta$  and  $x, y \in B$ .

**Proposition 10:** The relation  $R_\beta$  is a similar relation on the set of objects  $Ob$ .

Since  $\beta$  is a covering of  $Ob$ , then for any  $x \in Ob$  there exists  $B \in \beta$  such that  $x \in B$  hence  $x, x \in B \in \beta$  then  $x R_\beta x$ . Let  $x R_\beta y$  then there exists  $B \in \beta$  such that  $x, y \in B$  then  $y, x \in B$  hence  $y R_\beta x$ .

Proposition 10 for every  $x \in Ob$  we have:

$$R_\beta(x) = \bigcup_{B \in \beta(x)} B, \text{ where } \beta(x) = \{B \in \beta: x \in B\}$$

**Proof:**

$$y \in R_\beta(x) \Leftrightarrow \exists B \in \beta \text{ and } x, y \in B$$

$$\Leftrightarrow \exists B \in \beta \text{ and } x \in B \text{ and } y \in B$$

$$\Leftrightarrow \exists B \in \beta \text{ and } y \in B$$

$$\Leftrightarrow \bigcup_{B \in \beta(x)} B$$

Let  $\beta$  is the covering of  $Ob$ . Then we define the class  $\beta^* = \{R_\beta(x): x \in Ob\}$ .

**Proposition 11:** The class  $\beta^*$  is a symmetric covering of the set of objects  $Ob$  and  $R_\beta \subseteq R_{\beta^*}$ .

**Proof:**

- $x \in R_\beta(y) \Leftrightarrow \exists B \in \beta(y) \text{ and } x \in B \Leftrightarrow \exists B \in \beta(x) \text{ and } y \in B \Leftrightarrow y \in R_\beta(x)$
- Let  $(x,y) \in R_\beta \Rightarrow \exists B \in \beta \text{ and } x, y \in B \Rightarrow B \in \beta(x) \text{ and } B \in \beta(y) \Rightarrow B \in \beta(x) \cap \beta(y)$

$$\Rightarrow B \in \bigcup_{B \in \beta(x)} B = R_{\beta}(x) \in \beta^*$$

$$\Rightarrow x, y \in B \in \beta^*$$

$$\Rightarrow x, y \in R_{\beta^*}$$

Let  $A \subseteq Ob$  be any non empty subset of the set of objects. Then  $A$  is called a similar pre-class of  $R$  if for any  $x, y \in A \Rightarrow (x, y) \in R$ .

**Proposition 12** Every similar class  $R(x)$  is a maximal similar pre-class.

For an element  $x \in Ob$  we define a class called the pre-similar class of  $x$  as follows:

$L_R(x) = \{A \subseteq Ob: x \in A \text{ and } A \text{ is similar pre-class of } R\}$ .  
Let  $L_R = \{L_R(x): x \in Ob\}$  be the family of all pre-similar classes. Then we define a relation  $R^*$  on  $L_R$  by for any  $L_R(x), L_R(y) \in L_R$ ,  $L_R(x)R^*L_R(y)$  iff there exist  $A \in L_R(x)$  and  $B \in L_R(y)$  and  $A \cap B \neq \emptyset$ .

**Proposition 13:**

- The relation  $R^*$  on  $L_R$  is a similar relation
- $xRy$  iff  $L_R(x)R^*L_R(y)$  for any  $x, y \in Ob$

**Proof:**

- Since for any  $L_R(x) \in L_R$  and  $A \in L_R(x)$ ,  $A \cap A \neq \emptyset$  then  $L_R(x)R^*L_R(x)$  hence  $R^*$  is reflexive. Also if  $L_R(x)R^*L_R(y)$  then there exist  $A \in L_R(x)$  and  $B \in L_R(y)$  such that  $A \cap B \neq \emptyset$ , hence  $B \cap A \neq \emptyset$ , hence  $L_R(y)R^*L_R(x)$ . Then  $R^*$  is symmetric
- Firstly, we will prove that  $xRy \Rightarrow L_R(x)R^*L_R(y)$

Let  $(x, y) \in R \Rightarrow \{x, y\}$  is a similar pre-class of  $R$ .  
 $\Rightarrow$  there exist a similar class  $R(x)$  such that  $\{x, y\} \subseteq R(x)$  and  $R(x) \in L_R(x)$  but  $R$  is symmetric then  $R(x) \in L_R(y)$ , then there exist  $A = R(x) \in L_R(x)$  and  $B = R(x) \in L_R(y)$  and  $A \cap B \neq \emptyset$ , hence  $L_R(x)R^*L_R(y)$ .

Conversely, let for some  $x, y \in Ob$ ,  $L_R(x)R^*L_R(y)$  then there exist  $R(z) \in L_R(x)$  and  $R(z) \in L_R(y)$  a similar class of  $R$ . hence  $x \in R(z)$  and  $y \in R(z)$  then  $x, y \in R(z)$  hence  $xRy$ .

Let  $L_R(x)$  be the pre-similar class of  $x \in Ob$ . Then we define a set  $L_R^*(x) = \bigcup_{A \in L_R(x)} A$  called the  $R$ -link of  $x$ , where  $A \in L_R(x)$  and  $A \neq R(x)$ .

If  $L_R^*(x) = Ob$  then it is called open  $R$ - link of  $x$  and if  $L_R^*(x) \subset Ob$  then it is called closed  $R$ - link of  $x$ .

The class  $M = \{L_R^*(x): x \in Ob\}$  of all  $R$ - links of  $x \in Ob$  is a subbase of a topology on  $Ob$  called the linked topology and denoted  $\tau_{L_R^*}$ .

**Proposition 14:**

- For any  $x \in Ob$ ,  $L_R^*(x) \subseteq R(x)$
- The class  $M$  is a symmetric covering of  $Ob$

**Proof:**

- Let  $y \in L^*R(x) \Rightarrow y \in \bigcup_{A \in L_R(x)} A \Rightarrow$  there exists  $A = R(x) \in L_R(x)$  and  $y \in A$  then  $y \in R(x) \Rightarrow L^*R(x) \subseteq R(x)$
- For any  $x \in Ob$ ,  $x \in L_R^*(y) \Rightarrow M$  is covering of  $Ob$

Now let  $x \in L_R^*(y) \Rightarrow x \in \bigcup_{A \in L_R(x)} A$ :

- there exist  $A = R(x) \in L_R(y)$  and  $x \in R(x)$
- $x, y \in R(x)$
- $y \in L_R^*(x)$

then  $M$  is a symmetric covering of  $Ob$ .

**Proposition 15:**

- The linked topology  $\tau_{L_R^*}$  is finer than the similar topology  $\tau_R$ , where  $\tau_R$  is the topology generated by the subbase  $\{R(x): x \in Ob\}$
- $xRy \Rightarrow \exists$  open set  $u \in \tau_{L_R^*}$  and  $x, y \in u$

**Example:** Let  $Ob = \{c_1, c_2, \dots, c_7\}$  be the set of objects which is seven computers in a local network in a certain company. Let  $\tau$  be the irregular topology on the set of objects which induced by a general relation on  $Ob$  which makes the following graph. We define a similar relation  $R$  on the set of objects by: Two computers  $x$  and  $y$  are in relation by  $R$  iff the computer  $x$  has a copy of a certain program in the computer  $y$ .

Then we can define the similar classes of  $R$  as follows:

- $R(c_1) = \{c_1, c_2, c_4\}$ ,  $R(c_2) = \{c_1, c_2, c_3, c_4, c_5\}$ ,  $R(c_3) = \{c_2, c_3, c_5\}$ ,  $R(c_4) = \{c_1, c_2, c_5, c_6, c_4\}$ ,  $R(c_5) = \{c_2, c_3, c_4, c_6, c_5\}$ ,  $R(c_6) = \{c_4, c_5, c_6, c_7\}$ ,  $R(c_7) = \{c_6, c_7\}$ . Then we have  $L_R(c_1) = \{\{c_1\}, \{c_1, c_2\}, \{c_1, c_4\}, \{c_1, c_2, c_4\}\}$ ,  $L_R(c_2) = \{\{c_2\}, \{c_2, c_1\}, \{c_2, c_4\}, \{c_2, c_5\}, \{c_2, c_3\}, \{c_2, c_3, c_5\}, \{c_2, c_1, c_4\}, \{c_2, c_4, c_5\}, \{c_2, c_1, c_3, c_4, c_5\}\}$ ,  $L_R(c_3) = \{\{c_3\}, \{c_3, c_2\}, \{c_3, c_5\}, \{c_3, c_2, c_5\}\}$ ,  $L_R(c_4) = \{\{c_4\}, \{c_4, c_1\}, \{c_4, c_2\}, \{c_4, c_5\}, \{c_4, c_6\}, \{c_4, c_1, c_2\}, \{c_4, c_2, c_5\}, \{c_4, c_5, c_6\}, \{c_4, c_1, c_2, c_5, c_6\}\}$ ,  $L_R(c_5) = \{\{c_5\}, \{c_5, c_2\}, \{c_5, c_4\}, \{c_5, c_3\}, \{c_5, c_6\}, \{c_5, c_2, c_3\}, \{c_5, c_2, c_4\}, \{c_5, c_4, c_6\}, \{c_2, c_5, c_4, c_5, c_6\}\}$ ,  $L_R(c_6) = \{\{c_6\}, \{c_6,$

$c_4\}, \{c_6, c_5\}, \{c_6, c_7\}, \{c_6, c_4, c_7\}, \{c_6, c_5, c_7\}, \{c_6, c_4, c_5\}, \{c_4, c_5, c_6, c_7\}$

- $L_R(c_7) = \{\{c_7\}, \{c_6, c_7\}\}$ . Also, we have:  
 $L^*_R(c_1) = \{c_1, c_2, c_4\}$ ,  $L^*_R(c_2) = \{c_1, c_2, c_3, c_4, c_5\}$ ,  
 $L^*_R(c_3) = \{c_2, c_3, c_5\}$   $L^*_R(c_4) = \{c_1, c_2, c_5, c_6, c_4\}$ ,  
 $L^*_R(c_5) = \{c_2, c_3, c_4, c_5, c_6\}$ ,  $L^*_R(c_6) = \{c_4, c_5, c_6, c_7\}$ ,  $L^*_R(c_7) = \{c_7\}$

Then the linked topology  $\tau_{L^*_R}$  is finer than the similar topology, such that  $L^*_R(c_i) \subseteq R(c_i)$  for all  $i = 1, 2, \dots, 7$ .

For any subset A of the set of objects, we define two sets  $\underline{R}(A)$  and  $\overline{R}(A)$ , they are called the lower and upper similar classes of A by:

$$\underline{R}(A) = \bigcup \{R(x) : R(x) \subseteq A\}$$

and

$$\overline{R}(A) = \bigcup \{R(x) : R(x) \cap A \neq \emptyset\}$$

Let  $\tau_R$  be the topology induced by the subbase  $\{\underline{R}(A) : A \subseteq Ob\}$  this topology called the lower similar topology. Also we define the topology  $\overline{\tau}_R$  which is called the upper similar topology and generated by the subbase  $\{\overline{R}(A) : A \subseteq Ob\}$ .

**Proposition 16:** Let  $\tau_R$  and  $\overline{\tau}_R$  be the lower and upper similar topologies then:

- $\overline{\tau}_R \subseteq \tau_R$  if R is an equivalence relation
- $\tau_R \subseteq \overline{\tau}_R$  if R is a similar relation
- $\tau_R$  and  $\overline{\tau}_R$  are in general not comparable if R is a general relation

The following proposition present another way to generate topologies from similarity relations.

**Proposition 17:**  $\tau^{**}_R = \{A \subseteq Ob : \forall x \in A, R(x) \subseteq A\}$  is a topology on Ob.

**Proof:**

- $Ob, \emptyset \in \tau^{**}_R$  is clearly
- If  $A_1, A_2, \dots \in \tau^{**}_R$  and  $x \in \bigcup_i A_i$  for some  $i$ , then  $R(x) \subseteq A_i$  then  $R(x) \subseteq \bigcup_i A_i$  hence  $\bigcup_i A_i \in \tau^{**}_R$
- Let  $A_1, A_2 \in \tau^{**}_R$ , then  $\forall x \in A_1 \cap A_2$  we have  $R(x) \subseteq A_1$  and  $R(x) \subseteq A_2$  hence  $R(x) \subseteq A_1 \cap A_2$  then  $A_1 \cap A_2 \in \tau^{**}_R$

**Example:** Consider  $Ob = \{a, b, c, d\}$  be the set of objects with a similar relation R its similar classes are:

- $R(a) = \{a, c\}$ ,  $R(b) = \{b, d\}$ ,  $R(c) = \{a, c, d\}$  and  $R(d) = \{b, c, d\}$ . Then:  $\tau_R = \{Ob, \emptyset, \{c\}, \{d\}, \{c, d\}, \{a, c\}, \{b, d\}, \{a, c, d\}, \{b, c, d\}\}$ ,  $\overline{\tau}_R = \{Ob, \emptyset, \{d, c\}, \{a, c, d\}, \{b, c, d\}\}$  and  $\tau^{**}_R = \{Ob, \emptyset\}$  then  $\tau^{**}_R \subseteq \overline{\tau}_R \subseteq \tau_R$
- The conjugate relation  $\overline{R}$  of R is defined by  $(x, y) \in \overline{R}$  iff  $(x, y) \notin R$  or  $x = y$

**Proposition 18:**

- $\underline{R} \cap \overline{R} = I$ , I is the identity relation
- $\overline{R}$  is a similar relation
- $\overline{\overline{R}} = R$

**Proof:**

- $(x, y) \in R \cap \overline{R}$  iff  $x = y \Rightarrow R \cap \overline{R} = I$
- $(x, x) \in \overline{R}$  such that  $x = x$  and if  $(x, y) \in \overline{R}$  then  $(x, y) \notin R$  or  $x = y$  then  $(y, x) \notin R$  or  $y = x$  hence  $(y, x) \in \overline{R}$
- $(x, y) \in \overline{\overline{R}} \Leftrightarrow (x, y) \notin \overline{R}$  or  $x = y \Leftrightarrow (x, y) \in R$  or  $x = y \Leftrightarrow (x, y) \in R$

**Example:** Let  $Ob = \{a, b, c, d\}$  be the set of objects with the similar relation  $R = \{(a, a), (b, b), (c, c), (d, d), (d, c), (c, d), (d, b), (b, d), (c, b), (b, c), (b, a), (a, b)\}$ . Then  $\overline{R} = I \cup (Ob \times Ob - R) = \{(a, a), (b, b), (c, c), (d, d), (c, a), (a, c), (d, a), (a, d)\}$ .

**Topologies generated using dominance (pre-order) relations:** For a long time, many mathematicians believed that there is a large deviation between abstract topological structures and computing<sup>[12-14]</sup>.

A relation R on a set Ob is called a dominance relation (pre-order) whenever R is both reflexive and transitive. If x is related to y, we write  $xRy$  and say that x dominances y. The set  $R(y) = \{x : yRx\}$  is called the before set.

**Example:** Let  $Ob = \{1, 2, 3, 4, 5, 6\}$  and  $(x, y) \in R$  if and only if  $x \mid y$ ,  $x, y \in Ob$ .

$R = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (4, 4), (5, 5), (6, 6), (1, 6), (2, 2), (2, 4), (2, 6), (3, 3), (3, 6)\}$ . The relation R is a dominance relation.

In a finite space  $(Ob, \tau)$ , it is clear that  $\tau^c = \{Ob - G : G \in \tau\}$  is also a topology.

A subset  $F \subset X$  is a closed set iff  $F = \bigcup_{y \in F} Fy$  such that  $Fy = \{x : yRx\}$  ( $Fy$  is the smallest closed set about  $x$ ). This is the dual of our representation of open sets.

If  $R$  is a dominance relation on a set  $Ob$ , then its dual  $R_D$  is defined by the requirement  $yR_D x$  if and only if  $xRy$ .

A point  $x$  in a subset  $U$  of  $Ob$  is insulated from  $Ob - U$  if and only if there is no point  $y$  in  $Ob - U$  such that  $y$  dominates  $x$ .

**Lemma 19:** Let  $R$  be a dominance relation on a set  $Ob$ ,  $U \subset Ob$ ,  $P \in U$  the following are equivalent s:

- $P$  is insulated from  $Ob - U$
- $P \in U$ ,  $(y, P) \in R$ , then  $y \in U$

**Proof:** First, consider  $p \in U$  is insulated from  $Ob - U$ ,  $(y, p) \in R$ . Let  $y \notin U$ , then  $y \in Ob - U$ . So  $(y, p) \notin R$ , but  $(y, p) \in R$ , a contradiction, then  $y \in U$ .

Second, consider  $p \in U$ ,  $x \in Ob - U$ , suppose  $(x, p) \in R$ . Then  $x \in U$  contradicts that  $x \notin U$ , then  $\tau_R = \{U \subset Ob : x \text{ is insulated from } Ob - U, \forall x \in U\}$   $(x, p) \notin R$ .

**Proposition 20:** If  $R$  is a dominance relation on a set  $Ob$ , then is topology on  $Ob$ .

**Proof:** Clearly  $Ob$  and  $\phi$  are elements in  $\tau_R$  let  $U_i \in \tau_R$  for every  $i \in I$ . For any  $x \in \bigcup_{i \in I} U_i$  and  $(y, x) \in R$ , there is  $i_0 \in I$  such that  $x \in U_{i_0}$ . By openness of  $U_{i_0}$ , we have  $y \in U_{i_0} \subset \bigcup_{i \in I} U_i$ . Therefore  $\bigcup_{i \in I} U_i \in \tau_R$ . Also, if  $A$  and  $B$  are elements of  $\tau_R$ , then  $\tau_R = \{Ob, \phi, \{d\}\}$ .

According to the above proposition we give the following algorithm to check the openness of a subset  $U \subset Ob$  with respect to a dominance relation  $R$ .

**Openness algorithm:**

- i- Find  $Ob - U$
- ii- Investigates the existence of any pair  $(a, b) \in R$ ,  $a \in Ob - U$ ,  $b \in U$ , we have two cases:
  - If there exists such pair, then  $U$  is not open
  - If there is not such pair  $(a, b)$ , then  $U$  is open

The following example (Table 1) is an application for the above algorithm.

Table 1: Application for Openness algorithm

P(Ob)	Insulated points	Open sets
{1}	√	√
{2}		
{3}		
{4}		
{1, 2}	√	√
{1, 3}		
{1, 4}		
{2, 3}		
{2, 4}		
{3, 4}		
{1, 2, 3}	√	√
{1, 2, 4}		
{2, 3, 4}		
{1, 3, 4}		
Ob	√	√
$\phi$	√	√

Then the induced topology is  $\tau_R = \{Ob, \phi, \{1\}, \{1, 2\}, \{1, 2, 3\}\}$ .

Let  $Ob$  be the set of objects and let  $R$  be any binary relation on  $Ob$ . The relation  $R$  gives rise to a closure operator  $cl_R$  as follows:

$$cl_R(A) = AU\{y \in X \mid \exists x \in A : (y, x) \in R\} \text{ for every } A \subseteq Ob$$

**Lemma 21:** The interior operator corresponding to  $cl_R$  is given by:

$$int_R(A) = \{y \in A : \forall x \in A^c, \sim yRx\}$$

**Proof:**  $int_R(A) = [cl_R(A^c)]^c$   
 $= \{Ac \cup \{y \in X : \exists x \in A^c, (y, x) \in R\}\}^c$   
 $= A \cap \{y \in X : \exists x \in A^c, (y, x) \in R\}^c$   
 $= A \cap \{y \in X : \forall x \in A^c, (y, x) \notin R\}^c$   
 $= \{y \in A : \forall x \in A^c, \sim yRx\}$

Thus the interior operator of  $A$  consist of those elements of  $A$  which are not  $R$ -related to any elements outside  $A$ .

**Lemma 22:** For any relation  $R$  on  $Ob$ ,  $(Ob, cl_R)$  is closure space.

In the following we will give an example (Table 2) for closure space generated by a general relation.

**Example:** Consider  $Ob\{a, b, c\}$  and  $R$  is a binary relation on  $Ob$ ,  $R = \{(a, b), (c, b), (a, c)\}$ . Then we have the Table 2 for closures and interiors of the subsets of  $Ob$ : We note from Table 2 that:

- $cl_R(\phi) = \phi$
- $A \subseteq cl_R(A)$
- $cl_R(A \cup B) = cl_R(A) \cup cl_R(B)$  for all  $A, B \subseteq Ob$

Table 2: Closure space generated by a general relation

A	Cl <sub>R</sub> (A)	int <sub>r</sub> (A)
{a}	{a}	∅
{b}	{b, c}	{b}
{c}	{a, c}	∅
{a, b}	Ob	{b}
{a, c}	{a, c}	∅
{b, c}	Ob	{b, c}
Ob	Ob	Ob
∅	∅	∅

Table 3: Closures and interiors of the subsets of Ob

A	Cl <sub>R</sub> (A)	intr <sub>R</sub> (A)
{1}	{1}	{1}
{2}	{2, 3}	∅
{3}	{2, 3}	∅
{1, 2}	Ob	{1}
{1, 3}	Ob	{1}
{2, 3}	{2, 3}	{2, 3}
Ob	Ob	Ob
∅	∅	∅

**Lemma 23** If Ob be non-empty set and R is transitive relation, then (x, cl<sub>R</sub>) is topological space.

**Example:** Consider the relation R = {(1, 1), (2, 3), (3, 2), (2, 2)} on Ob = {1, 2, 3}. Table 3 shows closures and interiors of the subsets of Ob.

From Table 3, we have:

- cl<sub>R</sub>(∅) = ∅
- A ⊆ cl<sub>R</sub>(A)
- cl<sub>R</sub>(A ∪ B) = cl<sub>R</sub>(A) ∪ cl<sub>R</sub>(B) for all A, B ⊆ Ob
- cl<sub>R</sub>(cl<sub>R</sub>(A)) = cl<sub>R</sub>(A) for all A ⊆ Ob

**Topologies generated using general binary relations:**

The basic aim of this section is to generate topological structures using the lower and the upper approximations of any binary relation. Given general approximation space A = (Ob, R) where R here is any general binary relation on Ob. For any subset X of Ob we define lower and upper approximations as follows:

$$\underline{R}(X) = \{x \in Ob : \forall y((x, y) \in R \Rightarrow y \in X)\}$$

and

$$\overline{R}(X) = \{x \in Ob : \exists y((x, y) \in R \wedge y \in X)\}$$

Then the following structures are topologies on Ob:

$$\tau_{-1} = \{G \subseteq Ob : \underline{R}(G) = G\}$$

$$\tau_{-2} = \{G \subseteq Ob : \underline{R}^2(G) = \underline{R}(\underline{R}(G)) = G\}$$

$$\tau_{-3} = \{G \subseteq Ob : \underline{R}^3(G) = \underline{R}(\underline{R}(\underline{R}(G))) = G\} \dots$$

$$\tau_{-n-1} = \{G \subseteq Ob : \underline{R}^{n-1}(G) = G, n = |Ob|\}$$

These topologies have the property that:  $\tau_{-1} \subseteq \tau_{-2} \subseteq \dots \subseteq \tau_{-n-1}$ .

Also, if we deal with the upper approximation instead of the lower approximation we can construct the following topologies:

$$\tau_1 = \{G \subseteq Ob : \overline{R}(G) = G \vee \overline{R}(G) = \emptyset\}$$

$$\tau_2 = \{G \subseteq Ob : \overline{R}^2(G) = \overline{R}(\overline{R}(G)) = G \vee \overline{R}(\overline{R}(G)) = \emptyset\}$$

$$\tau_3 = \{G \subseteq Ob : \overline{R}^3(G) = \overline{R}(\overline{R}(\overline{R}(G))) = G \vee \overline{R}(\overline{R}(\overline{R}(G))) = \emptyset\} \dots$$

$$\tau_{n-1} = \{G \subseteq Ob : \overline{R}^{n-1}(G) = G \vee \overline{R}^{n-1}(G) = \emptyset, n = |Ob|\}$$

These topologies have the property that

$$\tau_1 \subseteq \tau_2 \subseteq \dots \subseteq \tau_{n-1}$$

In the following we will give some illustrative examples and remarks.

**Example:** Let Ob = {a, b, c, d} be the universe and let R = {(a, b), (c, d), (b, d), (d, a), (c, b)} be a general binary relation on Ob. Then we have the following topologies on Ob using the lower approximation:

$$\tau_{-1} = \{Ob, \emptyset, \{a, b, d\}\}$$

$$\tau_{-2} = \{Ob, \emptyset, \{a, b, d\}\}$$

$$\tau_{-3} = \{Ob, \emptyset, \{a\}, \{b\}, \{d\}, \{a, b\}, \{a, d\}, \{b, d\}, \{a, b, d\}\}$$

If we made more iteration to introduce more topologies using the lower approximation we will obtain that:  $\tau_{-4} = \tau_{-1}$ ,  $\tau_{-5} = \tau_{-2}$  and  $\tau_{-6} = \tau_{-3}$  and so on.

Also we have the following topologies on Ob using the upper approximation:

$$\tau_1 = \{Ob, \emptyset, \{c\}\}$$

$$\tau_2 = \{Ob, \emptyset, \{c\}\}$$

$$\tau_3 = \{Ob, \emptyset, \{d\}, \{c\}, \{a, c\}, \{b, c\}, \{a, b, c\}, \{a, c, d\}, \{b, c, d\}\}$$

If we made more iteration to introduce more topologies using the upper approximation we will obtain that:  $\tau_4 = \tau_1$ ,  $\tau_5 = \tau_2$  and  $\tau_6 = \tau_3$  and so on.

**Remark:** If the relation R on the universe Ob is constant, then all topologies induced by the lower or the upper approximations are indiscrete.

If the relation R on the universe Ob is identity or contain the identity relation, then all topologies induced by the lower or the upper approximations are discrete.

If we made more iteration to introduce more topologies using the lower approximation or the upper approximation, then all new iterations will introduce the same topologies we before obtained.

Another method for constructing topologies using the lower and the upper approximations is presented bellow:

All the following are topologies on Ob:

$$\begin{aligned} \tau_{-1} &= \{\underline{R}(G) : \underline{R}(G) = \bar{R}(G)\} \\ \tau_{-2} &= \{\underline{R}^2(G) : \underline{R}^2(G) = \bar{R}^2(G)\} \\ \tau_{-3} &= \{\underline{R}^3(G) : \underline{R}^3(G) = \bar{R}^3(G)\} \\ \dots, \tau_{-n-1} &= \{\underline{R}^{n-1}(G) : \underline{R}^{n-1}(G) = \bar{R}^{n-1}(G)\} \end{aligned}$$

Also, all the following structures are topologies on Ob.  $\tau^1 = \{\underline{R}(\bar{R}(G)) : \underline{R}(\bar{R}(G)) = \bar{R}(\underline{R}(G))\}$ ,  $\tau^2 = \{\underline{R}(\bar{R}(\underline{R}(G))) : \underline{R}(\bar{R}(\underline{R}(G))) = \bar{R}(\underline{R}(\bar{R}(G)))\}$  and so on.

**Example:** According to Example we have:

$$\begin{aligned} \tau_{-1} &= \{\text{Ob}, \varphi, \{d\}, \{a, b, c\}\} \\ \tau_{-2} &= \{\text{Ob}, \varphi, \{d\}\} \\ \tau_{-3} &= \{\text{Ob}, \varphi, \{d\}\} \end{aligned}$$

## RESULTS

The basic aim of our paper is to generate topological structures that are used to generate many topological measures and then using it in data reduction. First, we are used general binary relation to generate topological structures using the lower and the upper approximations. Make use of these methods we succeeded to introduce many topologies such as lower topologies from degree 1 to degree n where n is any order. Also, the topologies from degree 1 to n using upper approximations. Second, we are used the approach of closure operator and interior operator to induced topological structures.

We govern the importance of these topologies by the specified properties of relations that it are generate. Lemmas, propositions and theorems are mentioned to filter these topology to know which topology is used in applications of data reduction.

Table 4: The information system

Ob	a1	a2	a3	a4
x <sub>1</sub>	1	2	9	6
x <sub>2</sub>	3	2	6	2
x <sub>3</sub>	3	6	3	3
x <sub>4</sub>	4	2	2	3
x <sub>5</sub>	6	6	5	4

In our application example we succeeded to obtain reducts of multi-valued (set-valued) datasets that handled differently using ROSETTA software and proved that our solution is more accurate than solution in<sup>[10,14]</sup>.

Now if  $\tau_{I_{At}}$  is the topology induced by  $\{\text{Int}_{At}(X) : X \subseteq \text{Ob}\}$  ( $\tau_{C_{At}}$  or  $\tau_{N_{At}}$  can be used alternately), then when  $\tau_{i,j} = \tau_{I_{At}}$  the set  $\{a_i, a_j\}$  is a second order reduct of At in S. On the other hand, if  $\tau_{i,j} \neq \tau_{I_{At}}$  for all i, j = 1,2,...,n we must calculate the highest topologies  $\tau_{1,2,3}, \dots, \tau_{n-2,n-1,n}$  and the subset  $\{a_i, a_j, a_k\}$  is a third order reduct of At in S when  $\tau_{i,j,k} = \tau_{I_{AT}}$ . By the same manner, we can define a high order reducts of At in S.

In each case, the topological core of At in S is the intersection of all reducts (intersection of all the same order reducts). This core called the interior core and denoted  $\text{Core}_{\text{int}}(\text{At})$ . By the same terminology, we can define the closure core ( $\text{Core}_{C1}(\text{At})$ ) and the neighborhood core ( $\text{Core}_N(\text{At})$ ).

Illustrated Example Consider the information system shown by Table 4 and if we choose r = 2, then  $N_{a_i}(x, r) = \{y \in \text{Ob} : |f_{a_i}(x) - f_{a_i}(y)| \leq 2\}$ , hence we have the following subbases:

$$\begin{aligned} \zeta_1 &= \{\{x_1, x_2, x_3\}, \{x_1, x_2, x_3, x_4\}, \{x_2, x_3, x_4, x_5\}, \{x_4, x_5\}\} \\ \zeta_2 &= \{\{x_1, x_2, x_4\}, \{x_3, x_5\}\} \\ \zeta_3 &= \{\{x_1\}, \{x_3, x_4, x_5\}, \{x_2, x_5\}, \{x_3, x_4\}, \{x_2, x_3, x_5\}\} \\ \zeta_4 &= \{\{x_2, x_3, x_4, x_5\}, \{x_1, x_5\}, \text{Ob}\} \end{aligned}$$

The corresponding bases are:

$$\begin{aligned} \beta_1 &= \{\{x_1, x_2, x_3\}, \{x_1, x_2, x_3, x_4\}, \{x_2, x_3, x_4, x_5\}, \{x_4, x_5\}, \{x_4\}, \{x_2, x_3\}, \{x_2, x_3, x_4\}\} \\ \beta_2 &= \{\{x_1, x_2, x_4\}, \{x_3, x_5\}\} \\ \beta_3 &= \{\{x_1\}, \{x_3, x_4, x_5\}, \{x_2, x_5\}, \{x_3, x_4\}, \{x_2, x_3, x_5\}, \{x_5\}, \{x_3\}, \{x_3, x_5\}\} \\ \beta_4 &= \{\{x_2, x_3, x_4, x_5\}, \{x_1, x_5\}, \{x_5\}, \text{Ob}\} \end{aligned}$$

The corresponding topologies are:

$$\begin{aligned} \tau_1 &= \{\text{Ob}, \varphi, \{x_1, x_2, x_3\}, \{x_1, x_2, x_3, x_4\}, \{x_2, x_3, x_4, x_5\}, \{x_4, x_5\}, \{x_4\}, \{x_2, x_3\}, \{x_2, x_3, x_4\}\} \\ \tau_2 &= \{\text{Ob}, \varphi, \{x_1, x_2, x_4\}, \{x_3, x_5\}\} \end{aligned}$$



$$\tau_3 = \{Ob, \phi, \{x_1\}, \{x_3, x_4, x_5\}, \{x_2, x_5\}, \{x_3, x_4\}, \{x_2, x_3, x_5\}, \{x_5\}, \{x_3\}, \{x_3, x_5\}, \{x_1, x_2, x_5\}, \{x_1, x_3, x_4, x_5\}, \{x_1, x_2, x_3, x_5\}, \{x_1, x_3, x_4\}, \{x_1, x_5\}, \{x_1, x_3, x_5\}, \{x_1, x_3\}, \{x_2, x_3, x_4, x_5\}, \{x_2, x_3, x_5\}, \{x_3, x_4, x_5\}\}$$

$$\tau_4 = \{Ob, \phi, \{x_2, x_3, x_4, x_5\}, \{x_1, x_5\}, \{x_5\}\}$$

If we considered the set of all attributes then  $\tau_{N_{At}}$  is the discrete topology, but the second order topologies are given such that:  $\tau_{1,2} \neq \tau_{N_{At}}, \tau_{1,3} = \tau_{N_{At}}, \tau_{1,4} \neq \tau_{N_{At}}, \tau_{2,3} = \tau_{N_{At}}, \tau_{2,4} \neq \tau_{N_{At}}, \tau_{3,4} \neq \tau_{N_{At}}$ . Then  $\{a_1, a_3\}$  and  $\{a_2, a_3\}$  are second order reducts of At and the second order core is given by  $Core_N(At) = \{a_3\}$ .

**DISCUSSION**

In this section, we briefly describe the main idea of our work using the rheumatic fever datasets. No doubt that, the rheumatic fever is a very common disease. It has many symptoms differs from patient to another but though the diagnosis it is the same. So, we obtained the following data on seven rheumatic fever patients from Banha fever hospital, Egypt. All patients are between 9-12 years old with history of Arthurian began from age 3-5 years. This disease has many symptoms and it is usually started in young age and still with the patient along his life. Table 5 introduced the seven patients characterized by 8 symptoms (Attributes) using them to decide the diagnosis for each patient (Decision Attribute). Table 6 shows the rheumatic fever information system.

Table 5: Rheumatic fever data

Attribute name	Attribute values	Attribute refers to
Sex (S)	S <sub>1</sub>	Male
	S <sub>2</sub>	Female
Pharyngitis (P)	yes	Yes
	no	No
Arthritis (A)	a <sub>0</sub>	No arthritis
	a <sub>1</sub>	Began in the knee
	a <sub>2</sub>	Began in the ankle
Carditis (C)	r <sub>1</sub>	Affected
	r <sub>2</sub>	Not affected
Chorea (Ch)	yes	Yes
	no	No
ESR	e <sub>1</sub>	Normal
	e <sub>2</sub>	High
Abdominal pain (Ap)	p <sub>1</sub>	Absent
	p <sub>2</sub>	Present
Headache (H)	yes	Yes
	no	No
Diagnosis (D)	d <sub>1</sub>	Rheumatic arthritis
	d <sub>2</sub>	Rheumatic carditis
	d <sub>3</sub>	Rheumatic arthritis and carditis

Let us consider the topological space  $\tau_a$  generated using binary relation defined on the attribute a. Also, using the same terminology the topological space  $\tau_b$  is the topology generated using general relation defined on a subset of attributes B of all condition attributes At. The decision attribute generates the topology  $\tau_D$ .

Now, we will use the following suggestion, The set of attributes  $B \subseteq At$  is called a reduct if  $\tau_B \leq \tau_D$  and B is minimal, where:

$$(\tau_B \leq \tau_D \text{ iff } \forall G \in \tau_B, \exists G' \in \tau_D \text{ s.t. } G \subseteq G', G, G' \neq U)$$

The attribute  $a \in At$  is called the core if  $|\tau_a| > |\tau_b|, \forall a, b \in At, b \neq a$ .

When the classical technique of rough set theory (ROSETTA software)<sup>[4,5,13]</sup> used to obtain reducts and core of our data we found that we have 8 reducts of Table 6 with out any intersections among them. So, we do not have any core of Table 6. The set of obtained reducts is as follows:

$$\begin{aligned} Red(At) = & \{ \{S \vee C \vee Ch\}, \{S \vee A \vee C\}, \\ & \{S \vee P \vee A \vee Ap\}, \{P \vee A \vee C \vee K \vee Ap\}, \\ & \{C \vee ESR\}, \{P \vee A \vee ESR \vee Ap\}, \\ & \{A \vee C \vee H\}, \{P \vee A \vee Ap \vee H\} \end{aligned}$$

Now, after getting the reducts of Table 6 using the ROSETTA software. We will convert Table 5-8 using Table 7.

Now we will apply the above contributions on Table 7 where  $Ob = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$  is the set of objects, the set of condition attributes is  $At = \{\alpha, \beta, \delta\}$  and the decision attribute is the diagnosis D.

According to the binary relation  $R_B = \{(x, y), f_B(x)^c \subseteq f_B(y), B \subseteq At\}$  we can construct the following topologies:

$$\tau_\alpha = \{Ob, \phi, \{x_2\}, \{x_3\}, \{x_2, x_3\}\}, \tau_\beta = \{Ob, \phi\}, \tau_\delta = \{Ob, \phi\},$$

$$\tau_{\alpha\beta} = \{Ob, \phi\}, \tau_{\alpha\delta} = \{Ob, \phi\}, \tau_{\beta\delta} = \{Ob, \phi\}, \tau_{\alpha\beta\delta} = \{Ob, \phi\}$$

Table 6: Rheumatic fever information system

Attributes	S	P	A	C	Ch	ESR	Ap	H	D
<b>Patients</b>									
p1	S <sub>2</sub>	yes	a <sub>1</sub>	r <sub>1</sub>	yes	e <sub>1</sub>	p <sub>1</sub>	no	d <sub>3</sub>
p2	S <sub>1</sub>	yes	a <sub>1</sub>	r <sub>1</sub>	yes	e <sub>2</sub>	p <sub>1</sub>	yes	d <sub>3</sub>
p3	S <sub>2</sub>	yes	a <sub>2</sub>	r <sub>1</sub>	no	e <sub>1</sub>	p <sub>1</sub>	no	d <sub>3</sub>
p4	S <sub>1</sub>	yes	a <sub>1</sub>	r <sub>2</sub>	no	e <sub>1</sub>	p <sub>1</sub>	no	d <sub>1</sub>
p5	S <sub>1</sub>	no	a <sub>0</sub>	r <sub>1</sub>	no	e <sub>1</sub>	p <sub>2</sub>	no	d <sub>2</sub>
p6	S <sub>1</sub>	yes	a <sub>1</sub>	r <sub>1</sub>	no	e <sub>2</sub>	p <sub>1</sub>	no	d <sub>3</sub>
p7	S <sub>1</sub>	yes	a <sub>2</sub>	r <sub>1</sub>	no	e <sub>1</sub>	p <sub>1</sub>	yes	d <sub>3</sub>

Table 7: Convert table

Attribute symbol	Refers to?	Attribute values	Refers to?
$\alpha$	{S, CH}	$\alpha_1$	S takes $s_1$
		$\alpha_2$	K takes $k_1$
		$\alpha_3$	Each of {S,K}takes $\{s_2, k_2\}$
$\beta$	{P, A, ESR}	$\beta_1$	F takes $f_1$
		$\beta_2$	A takes $a_1$
		$\beta_3$	A takes $a_2$
		$\beta_4$	E takes $e_4$
$\delta$	{C, Ap, H}	$B_5$	Each of {F, A, E}takes $\{f_2, a_0, e\}$
		$\delta_1$	R takes $r_1$
		$\delta_2$	P takes $p_1$
		$\delta_3$	H takes $h_3$
		$\delta_4$	Each of {R,P,H} takes $\{r_2, p_2, h_2\}$
D	Diagnosis	$d_1$	Rheumatic arthritis
		$d_2$	Rheumatic carditis
		$d_3$	Rheumatic arthritis and carditis

Table 8: Multi-valued information system

U	$\alpha$	$\beta$	$\delta$	D
$x_1$	$\{\alpha_2\}$	$\{\beta_1, \beta_2, \beta_3\}$	$\{\delta_1, \delta_2\}$	$\{d_3\}$
$x_2$	$\{\alpha_1, \alpha_2\}$	$\{\beta_1, \beta_2\}$	$\{\delta_1, \delta_2, \delta_3\}$	$\{d_3\}$
$x_3$	$\{\alpha_3\}$	$\{\beta_1, \beta_3, \beta_4\}$	$\{\delta_1, \delta_2\}$	$\{d_3\}$
$x_4$	$\{\alpha_1\}$	$\{\beta_1, \beta_2, \beta_4\}$	$\{\delta_2\}$	$\{d_1\}$
$x_5$	$\{\alpha_1\}$	$\{\beta_4\}$	$\{\delta_1\}$	$\{d_2\}$
$x_6$	$\{\alpha_1\}$	$\{\beta_1, \beta_2\}$	$\{\delta_1, \delta_2\}$	$\{d_3\}$
$x_7$	$\{\alpha_1\}$	$\{\beta_1, \beta_3, \beta_4\}$	$\{\delta_1, \delta_2, \delta_3\}$	$\{d_3\}$

Now we will apply the relation  $R_D = \{(x, y), f(x) \subseteq f(y)\}$  to deal with the decision attribute D and we can construct the following topology:

$$\tau_D = \{Ob, \emptyset, \{x_1, x_2, x_3, x_6, x_7\}, \{x_1, x_2, x_3, x_4, x_6, x_7\}, \{x_1, x_2, x_3, x_5, x_6, x_7\}\}$$

We observe that,  $\tau_\alpha \leq \tau_D$ , this leads to from the above contributions that  $\{\alpha\}$  is the reduct and it is the core.

Then we can get the degree of dependency for each attribute as follows:

$$\text{For } a = \alpha, \text{ we get } \gamma(\alpha, D) = \frac{2}{7}, \text{ for } a = \beta, \text{ we get}$$

$\gamma(\beta, D) = 0$  and for  $a = \delta, \gamma(\delta, D) = 0$ . But if we get the degree of dependencies for the other attributes we will find that:

$$\begin{aligned} \gamma(\{\alpha, \beta\}, D) &= \gamma(\{\alpha, \delta\}, D) \\ &= \gamma(\{\beta, \delta\}, D) = \gamma(C, D) = 0 \end{aligned}$$

Thus, the set of attributes of equal highest degree of dependency is the reduct of our system. So we

conclude that  $\{\alpha\}$  is the reduct of our data using the topological method also,  $\{\alpha\}$  is the core of our system.

Now, we observe that the reduction that we got by using the GMIS is contained in the reduction that we got using the discernibility matrix and this clears for us that our method for getting the reduction is more precise than using the ROSETTA method. Because, the ROSETTA method can not apply on general binary relations.

**Topological reduction of single valued datasets:** By reduction we mean if we can remove some data from the data table given in our information system preserving its basic properties. To express this idea more precisely, let  $S = (Ob, At, \{V_a : a \in At\}, f_a)$  be an information system (numerical system). Let  $r$  be a positive real, for each object  $x \in Ob$  and for  $a \in At$ ,  $N_a(x, r)$  is the a-neighborhood of  $x$  and defined by:

$$N_a(x, r) = \{y \in Ob : |f_a(x) - f_a(y)| \leq r\}$$

For any subset B of At, the B- neighborhood of x is defined by:

$$N_B(x, r) = \{y \in Ob : |f_a(x) - f_a(y)| \leq r \quad \forall a \in B\}$$

For any subset X of Ob, we define two mappings  $Int, Cl : P(Ob) \rightarrow P(Ob)$  as follows:

$$\begin{aligned} Int_B(X) &= \{x \in Ob : N_a(x, r) \subseteq X, \forall a \in B\}, s \\ Cl_B(X) &= \{x \in Ob : N_a(x, r) \cap X \neq \emptyset, \forall a \in B\} \end{aligned}$$

The classes  $\{Int_B(X) : X \subseteq Ob, B \subseteq At\}, \{Cl_B(X) : X \subseteq Ob, B \subseteq At\}$  and  $\{N_B(x, r) : x \in Ob, B \subseteq At\}$  are subbases of a topological spaces denoted  $\tau_i, \tau_c$  and  $\tau_n$  respectively.

Now let  $At = \{a_1, a_2, \dots, a_n\}$  and let  $\tau_{i_{a_1}}, \tau_{i_{a_2}}, \dots, \tau_{i_{a_n}}, \tau_{c_{a_1}}, \tau_{c_{a_2}}, \dots, \tau_{c_{a_n}}$  and  $\tau_{N_{a_1}}, \tau_{N_{a_2}}, \dots, \tau_{N_{a_n}}$  be the topologies induced by the subbases  $\{Int_{a_1}(X) : X \subseteq Ob\}, \{Int_{a_2}(X) : X \subseteq Ob\}, \dots, \{Int_{a_n}(X) : X \subseteq Ob\}, \{Cl_{a_1}(X) : X \subseteq Ob\}, \{Cl_{a_2}(X) : X \subseteq Ob\}, \dots, \{Cl_{a_n}(X) : X \subseteq Ob\}$  and  $\{N_{a_1}(X) : X \subseteq Ob\}, \{N_{a_2}(X) : X \subseteq Ob\}, \dots, \{N_{a_n}(X) : X \subseteq Ob\}$ , respectively. These topologies called interior, closure and neighborhood topologies respectively.

One of the two attributes  $a_i, a_j, i \neq j$  is called interior-dispensable in At if  $\tau_{i_{a_i}} = \tau_{i_{a_j}}$ , otherwise,  $a_i$  or  $a_j$  is indispensable in At. Let  $\tau_{i_{1,2}}, \tau_{i_{1,3}}, \dots, \tau_{i_{n-1,n}}$  be the

topologies induced by  $\tau_{I_{a_1}} \cup \tau_{I_{a_2}}, \tau_{I_{a_1}} \cup \tau_{I_{a_3}}, \dots, \tau_{I_{a_{n-1}}} \cup \tau_{I_{a_n}}$  if interior topologies are used (the same terminology used if closure topologies or neighborhood topologies is replaced).

### CONCLUSION

There are many approaches for obtaining topologies by relations and we used some of them in data reduction. These approaches were generalizations to Pawlak approaches namely, we ignored the notion of equivalence relations. Also, these approaches open the way for other approximations if we use the general topological recent concepts such as pre-open sets or semi-open sets. Make use of this terminology to obtain the missing values in incomplete datasets will be a good future work<sup>[1,4,5,6,16]</sup>. Implementing software for large data sets reduction using advanced programming languages will be also a good future work.

### REFERENCES

1. Brtko, V., E. Stokic and B. Srdic, 2008. Automated extraction of decision rules for leptin dynamics-A rough sets approach. *J. Biomed. Inform.*, 41: 667-674 DOI: 10.1016/j.jbi.2008.01.005
2. Davvaz, B., 2008. A short note on algebraic T-rough sets. *Inform. Sci.*, 178: 3247-3252. DOI: 10.1016/j.ins.2008.03.014
3. Degang, C., Y. Wenxia and L. Fachao, 2008. Measures of general fuzzy rough sets on a probabilistic space. *Inform. Sci.*, 178: 3177-3187. DOI: 10.1016/j.ins.2008.03.020
4. Hu, Q., D. Yu, J. Liu and C. Wu, 2008. Neighborhood rough set based heterogeneous feature subset selection. *Inform. Sci.*, 178: 3577-3594. DOI: 10.1016/j.ins.2008.05.024
5. Li, T., Y. Leung and W. Zhang, 2008. Generalized fuzzy rough approximation operators based on fuzzy coverings. *Int. J. Approx. Reason.*, 48: 836-856. DOI: 10.1016/j.ijar.2008.01.006
6. Liu, G., 2008. Axiomatic systems for rough sets and fuzzy rough sets. *Int. J. Approx. Reason.*, 48: 857-867. DOI: 10.1016/j.ijar.2008.02.001
7. Fotea, V.L., 2008. The lower and upper approximations in a hypergroup. *Inform. Sci.*, 178: 3605-3615. DOI: 10.1016/j.ins.2008.05.009
8. Mi, J., Y. Leung, H. Zhao and T. Feng, 2008. Generalized fuzzy rough sets determined by a triangular norm. *Inform. Sci.*, 178: 3203-3213. DOI: 10.1016/j.ins.2008.03.013
9. Min, F., Q. Liu and C. Fang, 2008. Rough sets approach to symbolic value partition. *Int. J. Approx. Reason.*, 49: 689-700. DOI: 10.1016/j.ijar.2008.07.002
10. Pattaraintakorn, P. and N. Cercone, 2008. A foundation of rough sets theoretical and computational hybrid intelligent system for survival analysis. *Comput. Math. Appl.*, 56: 1699-1708. DOI: 10.1016/j.camwa.2008.04.030
11. Qiana, Y., J. Liang and C. Dang, 2008. Consistency measure, inclusion degree and fuzzy measure in decision tables. *Fuzzy Sets Syst.*, 159: 2353-2377. DOI: 10.1016/j.fss.2007.12.016
12. Wang, X., J. Zhai and S. Lu, 2008. Induction of multiple fuzzy decision trees based on rough set technique. *Inform. Sci.*, 178: 3188-3202. DOI: 10.1016/j.ins.2008.03.021
13. Yang, Y. and R.I. John, 2008. Generalizations of roughness bounds in rough set operations. *Int. J. Approx. Reason.*, 48: 868-878. DOI: 10.1016/j.ijar.2008.02.002
14. Yao, Y. and Y. Zhao, 2008. Attribute reduction in decision-theoretic rough set models. *Inform. Sci.*, 178: 3356-3373. DOI: 10.1016/j.ins.2008.05.010
15. Zhao, S.E. and C.C. Tsang, 2008. On fuzzy approximation operators in attribute reduction with fuzzy rough sets. *Inform. Sci.*, 178: 3163-3176. DOI: 10.1016/j.ins.2008.03.022
16. Zhang, S., 2009. Mining class-bridge rules based on rough sets. *Expert Syst. Appl.*, 36: 6453-6460. DOI: 10.1016/j.eswa.2008.07.044