

# Membership function based on similarity and dissimilarity

E.Elsodany<sup>1</sup> and A.M.Kozae<sup>2</sup>

<sup>1</sup> Department of Physics and Engineering Mathematics, Faculty of Engineering, Tanta University, Tanta, Egypt

<sup>2</sup> Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt

## Abstract

Deciding membership function of an object to uncertain class of objects in an information system is essential particularly, all things considered, displaying issues. This paper will present a comparison between different ways of constructing membership function of an object. We initiate a concept based on dissimilarity and similarity between objects using values of their features. The fundamental tool is classes resulted from dissimilarity and similarity. We will also express a classification relation and minimum and maximum memberships correspond to these classifications.

**Keywords:** rough membership, dissimilarity measures and similarity.

## 1. Introduction

Organizations need information to analyze problems, control their work, make decisions and produce new services. The information system handles and processes information which can be considered the backbone of organizations. Information systems are in every branch of our daily life such as education, finance, healthcare, businesses, and government. A tool to deal with the uncertainty of these systems became an utmost necessity. Rough set theory was presented by Pawlak [1-3] in 1980's to deal with uncertain data or incomplete knowledge. Rough sets handled the uncertainty of data depending on approximations and equivalence relations. Another issue of classification is clustering which has many methods and the more similar objects in the same cluster than the dissimilar objects in a different cluster under a specific condition. The nature of collected data determines the measuring criteria of similarity. Algorithms of quantitative and qualitative data, graph diagrams, and similarity and dissimilarity measures were presented [4-7]. Membership of an object to a set is an important issue to identify participation of each object in this set. Membership was discussed [8] from a topological view which was extended from the definition of rough membership function[9]. Topological space has an intersection between basis but rough set classes have no intersection. [10] discussed the opinion of classes intersection in rough set and introduced a membership of an object to a class. In this case, an object can belong to more than one class, so it has different memberships minimum, maximum and average. From this point of view, we will introduce a membership function of an object to a set depending on the dissimilarity measures and similarity between objects. We will also define relations or classes then determine the membership of an object to a set. Comparison between minimum, maximum and membership of topological view will be mentioned. The decision of most, least membership value can be chosen by the expert.

The next sections of this paper is developed to preliminaries used in the paper, section two presents main definitions, results and examples. Section three is reserved for the more results and illustrations.

## 2. Preliminaries

In this section, we will mention some essential basics and definitions of rough set and information system representation. The dissimilarity and similarity between attributes of objects esteem are presented. Likewise, the membership of rough set and topology is also mentioned.

### Definition 2.1 [1]

The information system is a pair  $S = (U, A)$ , where  $U$  is a nonempty finite set called the universe and  $A$  is a nonempty finite set of attributes. Every attribute  $a \in A$  is a total function  $a: U \rightarrow V_a$ , where  $V_a$  is the set of values of  $a$  called the domain of  $a$ .

### Definition 2.2 [1]

Let  $B$  be a subset of attributes  $A$  ( $B \subset A$ ), a binary relation  $IND(B)$  called indiscernibility relation is defined as follows:

$$IND(B) = \{(x, y) \in U^2 : \forall a \in B, a(x) = a(y)\}$$

So  $IND(B)$  is an equivalence relation and  $IND(B) = \bigcap_{a \in B} IND(a)$

This view classify the objects depends on the equal features between objects as a result of this there is no intersection between relations.

### Definition 2.3[7]

Let an information system  $S = (U, A)$  and two objects  $x, y \in U$  described by  $a_k$  attributes subset of  $A$ . Mismatches of the attributes between two objects are the dissimilarity measures which is defined as

$$d(x, y) = \sum_{i=1}^k \delta(a_i(x), a_i(y))$$

Where

$$\delta(a_i(x), a_i(y)) = \begin{cases} 0 & : a_i(x) = a_i(y) \\ 1 & : a_i(x) \neq a_i(y) \end{cases}$$

This can be considered as another view of an information system classification. The produced classes are relations of dissimilarity between objects. The similarity between objects is the opposite way of dissimilarity or simply the number of matches between two objects.

**Definition 2.4** [9]

Let  $S = (U, A)$  be an information system and let  $\emptyset \neq X \subset U$ . The rough  $S$ -membership function of the set  $X$ , denoted by  $\mu_X^S$  is defined as follows

$$\mu_X^S(x) = \frac{|[x]_A \cap X|}{|[x]_A|} \text{ for } x \in U$$

**Definition 2.5** [8]

Let  $\tau$  be a topology on a finite set  $U$ , where its base is  $\beta$ , then the rough membership function is

$$\mu_X^\tau(x) = \frac{|(\cap \beta_x) \cap X|}{|\cap \beta_x|}, \beta_x \in \beta, x \in U$$

We will call this definition rough-topology membership function in the following section.

**3. Results**

In this section, we will illustrate the previous definitions and introduce a comparison between membership views. We will also work on the resulted relations of the dissimilarity and similarity relations.

**Definition 3.1**

Let an information system  $S = (U, A)$  and two objects  $x, y \in U$  are described by  $a_k$  attributes subset of  $A$ . The matches of the attributes between two objects is the similarity between objects which is defined as

$$s(x, y) = \sum_{i=1}^k \lambda(a_i(x), a_i(y))$$

Where

$$\lambda(a_i(x), a_i(y)) = \begin{cases} 0 & : a_i(x) \neq a_i(y) \\ 1 & : a_i(x) = a_i(y) \end{cases}$$

**Definition 3.2**

Let  $S = (U, A)$  be an information system,  $x_i$  and  $x_j$  subset of objects and  $x_i R x_j$  represents the relation between them which is resulted from the dissimilarity measures and similarity representation can be defined as

$$x_i R = \{x_j : x_i R x_j\} \text{ where } d(x_i, x_j) \neq 0 \text{ and } s(x_i, x_j) \neq 0$$

These relations are considered a classification of objects with respect to dissimilarity and similarity between objects. We will use these classes instead of the classical classes of an indiscernibility relation.

### Example 3.1

The following information system in Table 1 consists of 5-objects  $\{x_1, x_2, x_3, x_4, x_5\}$  and 3-attributes  $\{a_1, a_2, a_3\}$ .

U	attributes		
	$a_1$	$a_2$	$a_3$
$x_1$	1	5	20
$x_2$	1	9	10
$x_3$	1	5	10
$x_4$	4	8	15
$x_5$	3	5	20

Table 1 an information system

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$x_1$	0	2	1	3	1
$x_2$	2	0	1	3	3
$x_3$	1	1	0	3	2
$x_4$	3	3	3	0	3
$x_5$	1	3	2	3	0

Table 2 dissimilarity measures representation

Table 2 represents the dissimilarity measures of an information system of Table 1 and we apply Definition 2.3. This table represents the number of different attributes which can be considered dissimilarity representation of an information system. Each value in this table represents the number of the corresponding attributes of objects are far away from each others. Each value of Table 2 takes 4-values 0,1,2 and 3 because we have only 3-attributes. The value 0 means the corresponding 2-objects have an identical value for all attributes. The value 1 means only 1-mismatch attribute between the 2-objects. The value 2 means 2-mismatches between attributes but, value 3 means there is no match at all between objects.

It's obvious that similarity between objects is the complement of dissimilarity measures. Similarity can be obtained using Definition 3.1 to take 1 if the two objects have the same value and vice versa.

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$x_1$	3	1	2	0	2
$x_2$	1	3	2	0	0
$x_3$	2	2	3	0	1
$x_4$	0	0	0	3	0
$x_5$	2	0	1	0	3

Table 3 similarity representation

Table 3 shown represents the similarities between objects of table 1. We can observe that this table is the opposite of Table 2. The number of matches between objects is calculated instead of the number of mismatches. So the value 3 means the two objects are identical, but value 0 means there are no matches between objects. The value 1 and 2 mean there is one-match and two-match between objects, respectively.

The resulting classes of Table 2 using Definition 3.2 are

$$x_1 R = \{x_2, x_3, x_4, x_5\}, x_2 R = \{x_1, x_3, x_4, x_5\}, x_3 R = \{x_1, x_2, x_4, x_5\}$$

$$x_4 R = \{x_1, x_2, x_3, x_4\}, x_5 R = \{x_1, x_2, x_3, x_4\}$$

And classes of Table 3 are

$$x_1 R = \{x_1, x_2, x_3, x_5\}, x_2 R = \{x_1, x_2, x_5\}, x_3 R = \{x_1, x_2, x_3, x_5\}$$

$$x_4 R = \{x_4\}, x_5 R = \{x_1, x_3, x_5\}$$

We also observed the intersection between classes of dissimilarity and similarity alike classes of an indiscernibility relation.

Membership function is an important concept which identifies the participation of an object with respect to a set. From the above relations, we can observe that an object participates in more than one class so it can have more than one membership value with respect to the class. In the following, we will calculate rough membership of Definition 2.4 according to the classes of dissimilarity and similarity representation, respectively.

To calculate the membership, we should choose a set  $X$  so we will choose random sets, then we will apply Definition 2.4.

**Case 1:  $X = \{x_1, x_2\}$**

objects	$\mu_X^S(x_i)$ with respect to $x_i R$				
	$x_1 R$	$x_2 R$	$x_3 R$	$x_4 R$	$x_5 R$
$x_1$	—	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$x_2$	$\frac{1}{4}$	—	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$x_3$	$\frac{1}{4}$	$\frac{1}{4}$	—	$\frac{1}{2}$	$\frac{1}{2}$
$x_4$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	—	$\frac{1}{2}$
$x_5$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	—

Table 4 membership of case 1 dissimilarity representation

**Case 2:**  $X = \{x_1, x_2, x_3\}$

objects	$\mu_X^S(x_i)$ with respect to $x_i R$				
	$x_1 R$	$x_2 R$	$x_3 R$	$x_4 R$	$x_5 R$
$x_1$	—	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{3}{4}$
$x_2$	$\frac{1}{2}$	—	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{3}{4}$
$x_3$	$\frac{1}{2}$	$\frac{1}{2}$	—	$\frac{3}{4}$	$\frac{3}{4}$
$x_4$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	—	$\frac{3}{4}$
$x_5$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{4}$	—

Table 5 membership of case 2 dissimilarity representation

**Case 3:**  $X = \{x_1, x_2, x_3, x_4\}$

objects	$\mu_X^S(x_i)$ with respect to $x_i R$				
	$x_1 R$	$x_2 R$	$x_3 R$	$x_4 R$	$x_5 R$
$x_1$	—	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	1
$x_2$	$\frac{3}{4}$	—	$\frac{3}{4}$	$\frac{3}{4}$	1
$x_3$	$\frac{3}{4}$	$\frac{3}{4}$	—	$\frac{3}{4}$	1
$x_4$	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	—	1
$x_5$	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	—

Table 6 membership of case 3 dissimilarity representation

From Table 4 to Table 6, we can say that each object has a membership to  $X$  with respect to the corresponding  $x_i R$  which contains the object. And membership value of all objects belongs to  $x_i R$  with respect to  $X$  are equal. We need to have only one membership value of an object so we can choose minimum or maximum membership. In the following, we will apply Definition 2.5 in each case and compare the result with our previous results considering the minimum and maximum membership.

objects	Rough membership		Rough-topology membership
	Min	Max	
$x_1$	$\frac{1}{4}$	$\frac{1}{2}$	1
$x_2$	$\frac{1}{4}$	$\frac{1}{2}$	1

$x_3$	$\frac{1}{4}$	$\frac{1}{2}$	0
$x_4$	$\frac{1}{4}$	$\frac{1}{2}$	0
$x_5$	$\frac{1}{4}$	$\frac{1}{2}$	0

Table 7 dissimilarity representation membership comparison of case 1

objects	Rough membership		Rough-topology membership
	Min	Max	
$x_1$	$\frac{1}{2}$	$\frac{3}{4}$	1
$x_2$	$\frac{1}{2}$	$\frac{3}{4}$	1
$x_3$	$\frac{1}{2}$	$\frac{3}{4}$	1
$x_4$	$\frac{1}{2}$	$\frac{3}{4}$	0
$x_5$	$\frac{1}{2}$	$\frac{3}{4}$	0

Table 8 dissimilarity representation membership comparison of case 2

objects	Rough membership		Rough-topology membership
	Min	Max	
$x_1$	$\frac{3}{4}$	1	1
$x_2$	$\frac{3}{4}$	1	1
$x_3$	$\frac{3}{4}$	1	1
$x_4$	$\frac{3}{4}$	1	1
$x_5$	$\frac{3}{4}$	1	0

Table 9 dissimilarity representation membership comparison of case 3

From Table 7 to Table 9 we illustrate the comparison in each case and topological view of classes taking into account the intersection between them. But in our case resulting classes only intersect on one object. As a result of this, the membership is 1 and 0 according to the object belongs or not belongs to  $X$ , respectively. This makes the membership like the case of the classical crisp set. In some cases the maximum rough membership value equal to the topological rough membership value. This gives a chance to choose between more than membership value according to the problem we have.

In the following, we will also make a comparison of each case, but we will take into account similarity between objects and their corresponding classes. Tables will consider

the resulting minimum and maximum of rough membership then we will compare them with the topological view of rough membership.

objects	Rough membership		Rough-topology membership
	Min	Max	
$x_1$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$
$x_2$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$
$x_3$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$
$x_4$	0	0	0
$x_5$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$

Table 10 similarity representation membership comparison of case 1

objects	Rough membership		Rough-topology membership
	Min	Max	
$x_1$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{1}{2}$
$x_2$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{2}{3}$
$x_3$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{2}{3}$
$x_4$	0	0	0
$x_5$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{1}{2}$

Table 11 similarity representation membership comparison of case 2

objects	Rough membership		Rough-topology membership
	Min	Max	
$x_1$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{2}$
$x_2$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$x_3$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{2}{3}$
$x_4$	1	1	1
$x_5$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{2}$

Table 12 similarity representation membership comparison of case 3

We can conclude from the above comparison that in some cases, there is a match among minimum or maximum membership and topological membership values. In other cases, the value of topological membership is 0 but minimum and maximum values are not and



the opposite is true. This gives us a variety of choice and we can take a decision according to our desire.

## Conclusion

The proposed strategy for developing another view of membership function based on dissimilarity and similarity can help in decision making of uncertain issues. This view is constantly hard to be developed so our technique can take care of new issues of areas in social, financial aspects, arithmetic, and so on.

## References

- [1] Z. Pawlak, "Rough sets," *International Journal of Computer & Information Sciences*, vol. 11, pp. 341-356, 1982.
- [2] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*: Kluwer Academic Publishers, 1992.
- [3] Z. Pawlak and W. Marek, "Rough sets and information systems," *ICS. PAS. Reports*, vol. 441, pp. 481-485, 1981.
- [4] T. Alharthi and M. Elsafty, "Discovery of Similarity and Dissimilarity," *Journal of Mathematics Research*, vol. 8, p. 105, 2016.
- [5] S. Coppock and L. Mazlack, "Rough sets used in the measurement of similarity of mixed mode data," in *Fuzzy Information Processing Society, 2003. NAFIPS 2003. 22nd International Conference of the North American*, 2003, pp. 197-201.
- [6] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*: Elsevier, 2011.
- [7] Z. Huang, "A fast clustering algorithm to cluster very large categorical data sets in data mining," *DMKD*, vol. 3, pp. 34-39, 1997.
- [8] E. F. Lashin, A. M. Kozae, A. A. A. Khadra, and T. Medhat, "Rough set theory for topological spaces," *Int. J. Approx. Reasoning*, vol. 40, pp. 35-43, 2005.
- [9] Z. Pawlak and A. Skowron, "Rough membership functions," in *Advances in the Dempster-Shafer theory of evidence*, 1994, pp. 251-271.
- [10] Y. Yao and J. Zhang, "Interpreting fuzzy membership functions in the theory of rough sets," in *International Conference on Rough Sets and Current Trends in Computing*, 2000, pp. 82-89.