The Third AFSS(1998). 278－283

An Interval Approach to Fuzzy Pattern Recognition

Fatiha KARBOU(*), Fatima KARBOU(**), M. KARBOU(***)
(*):ENSEM, Dept. E.G, B. P. 8118, Oasis, Route d'El Jadida, Casablanca, Morocco
(**) IAV HASSAN II, dept. Topography, B.P. 6202, Rabat-instituts, Rabat, Morocco
(***): Sophia Concept, 47, rue Planquette Belvédère, Casablanca , Morocco

Abstract
The interval approach to the linguistic expression coding nears us to the human idea. Thus, what seems "weak" for a person can appear very weak for another person or for the same person in others circumstances. However, the utilization of intervals is not restrained to the cases of linguistic expression coding. Indeed, the interval can facilitate the solution of several problems.

Keywords: Fussy sets, inclusion degree, interval, codification of linguistic expressions.

1. Introduction

The use of linguistic expression such "small", "medium", "large", etc. to describe a given situation, near us to the human perception. Nevertheless, the attribution of these expressions can vary from one person to another. Consequently, the numerical coding of these expressions is very difficult

The interval approach [1] to the linguistic expression coding nears us to the human idea. Thus, what seems "weak" for a person can appear very weak for another person or for the same person in others circumstances. Therefore the attribution of linguistic expression depends on several parameters like the person itself and its environment. Therefore its value is displayed on an interval.

Generally, a linguistic expression attributed to a given situation is represented by fuzzy set characterized by a membership function. The α level of this function is an interval (see figure (1)). Therefore the notion of interval is implicit in the definition of fuzzy set relative to the linguistic expression.

Thus, the coding by interval is justified.

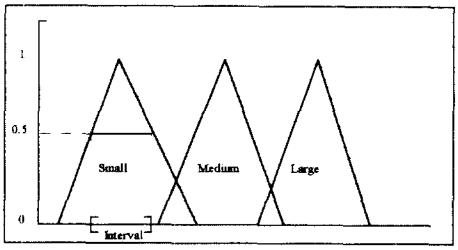

Fig.1 Example of memberships functions for the linguistic expressions

However, the utilization of intervals is not restrained to the cases of linguistic expression coding. Indeed, the interval can facilitate the solution of several problems. In this paper, we present a general form of fuzzy inclusion degree of two intervals, and, some interesting applications.

2. The fuzzy inclusion degree

The fuzzy inclusion degree inclusion introduced in [1,2] is the consequence of coding by interval. It measures the (mutual) inclusion degree of one interval in another. This degree depends on the distance D_{int} (I_1, I_2) that measures the dissimilarity between the intervals I_1 and I_2. It is a combination of three distances:

- A distance that measure the remoteness of the centers,
- A distance that measure the degree of the width of the intersection,
- A distance that measures the width of the difference of the sizes.

In this paragraph, we propose to study the influence of each of these three distances on the curve of fuzzy inclusion degree.

2.1. The role of the three functions

The use of the three functions seems to be heuristic. However, we must use all of them. In fact:

1. If we use only a distance that measures the remoteness of the centers (equation 2.4), then all of the concentric intervals will have the same inclusion degree as illustrated in the figure (2.a).
2. If we add just a distance that measures the degree of the width of the intersection (equation 2.5), then all of the concentric intervals with a very large difference of size will have the same fuzzy inclusion degree as illustrated in the figure (2.b).
3. If we add a distance that measures the width of the difference of the sizes (equations 2.6 and 2.7). The

result is satisfying as illustrated in the figures (2.c, 2.d).

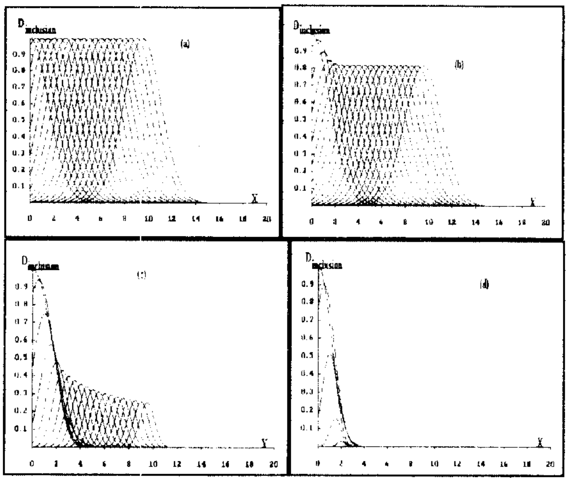

Fig. 2 The influence of the defined functions:
- (a): we use only the equation (2.4)
- (b): we use (2.4) and (2.5)
- (c): we use (2.4), (2.5) and (2.6)
- (d): we use (2.4), (2.5) and (2.7)

2.2. Study of the distances

Let $I_1=[a, b]$ and $I_2=[x, y]$, the intervals of description. Let $L_1= (b- a)$, the width of I_1, and $L_2= (y-x)$, the one of I_2, O_1 the center of I_1, and O_2 the one of I_2.

We are going to study the following equations:

$$D_{inclusion}(I_1, I_2)=\frac{1}{1+D_{int}(I_1,I_2)^k} \quad (2.1)$$

or

$$D_{inclusion}(I_1, I_2)=e^{-k\,D_{int}(I_1,I_2)} \quad (2.2)$$

With: $D_{int}(I_1, I_2)$ is the distance that measures the dissimilarity between the intervals. It is a combination of the following three distances:

$$D_{int}(I_1,I_2)=(f_L(L_1,L_2))^{k1}\left((f_d(L_1,L_2))^{k2}+(d(O_1,O_2))^{k3}\right)$$
(2.3)

Where:

$$d(O_1,O_2)=(O_1-O_2)^2 \quad (2.4)$$

$$f_d(L_1,L_2)=1-e^{-k(L_1-L_2)^2} \quad (2.5)$$

$$f_L(L_1,L_2)=1+\log(1+(L_1-L_2)^2) \quad (2.6)$$

$$=(1+(L_1-L_2)^2) \quad (2.7)$$

With $k1>0$, $k2>0$ and $k3>0$ are the coefficients of weighting. Their role consists of accenting or concealing the role of the corresponding functions.

We can say that:

(2. 4): measures the remoteness of the centers,

(2. 5): measure the degree of the width of the intersection,

(2. 6), (2.7): measures the width of the difference of the sizes.

We present some concentric intervals with different sizes. We have the reference interval $[a, b]=[10,12]$ centered at the beginning in 11. And $[x-1, x+1]$ the arbitrary intervals, when x varies from 1 to 19. Thus, the reference interval gets larger for the i^{rd} iterations when a becomes $a-(i-1)$ and b becomes $b+(i-1)$.

In the following sections, we illustrate the influence of k, k1, k2 and k3, on the curve of fuzzy inclusion degree using the equation (2.2). We use the same examples.

2.2.1. Classic case (k1 = k2 = k3 = 1)

In this case, the three functions have the same coefficients of weighting. We see that the fuzzy inclusion degree of I_1 in I_2 gets as greater as I_1 becomes closer to I_2. Its maximal value (≤ 1) is reached when the centers O1 and O2 of the two intervals coincide. It is equal to 1 when the two intervals are exactly identical, and more the difference of the sizes is important, more this value decreases (see figure 3).

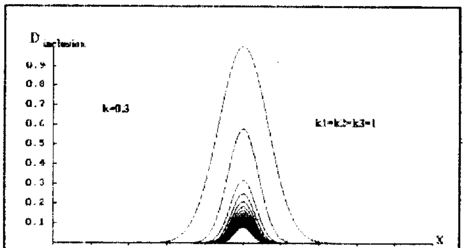

Fig. 3 The classic case: the three functions have the same pondering coefficients

2.2.2. The influence of k1

In this part, we fix k, k2, k3 and we vary k1, in order to illustrate the influence of the function (2.6) on the degree of inclusion.

We obtain the same results like the previous case (classic case). Indeed, the fuzzy inclusion degree of I_1

in I_2 gets as greater as I_1 becomes closer to I_2. Its maximal value (≤ 1) is reached when the centers O1 and O2 of the two intervals coincide. It is equal to 1 when the two intervals are exactly identical, and more the difference of the sizes is important, more this value is decreases.

However, we remark that:

The more k1 is decreasing, the more the level of bringing together of the inclusion curves is greater. Thus the maximal values of these curves are relatively brought together. Consequently, the number of curves no null for intervals of different sizes is elevated. (See figure 4)

As k1 is increasing, the curves are more distant and tend to get null when the two intervals have very different sizes. (See the figure 4)

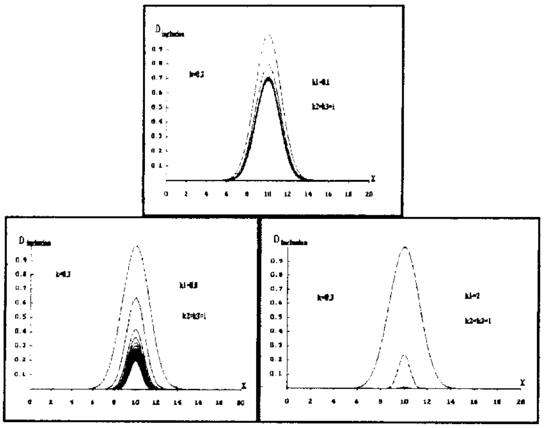

Fig. 4 The influence of k1

2.2.3. *The influence of k2*

In this part, we fix k, k1, k3 and we make vary k2, in order to illustrate the influence of the function (2.5) on the degree of inclusion.

Like the previous cases, the fuzzy inclusion degree of I_1 in I_2 gets as greater as I_1 becomes closer to I_2. Its maximal value (≤ 1) is reached when the centers O_1 and O_2 of the two intervals coincide. It is equal to 1 when the two intervals are exactly identical, and more the difference of the sizes is important, more this value is decreases.

We remark that:

The more k2 increases, the more the second curve tends to bring together of the first one. The other curves seem not influenced. This is owing to the fact that the exponential function is very sensible to the important difference of interval's size. (See figure 5)

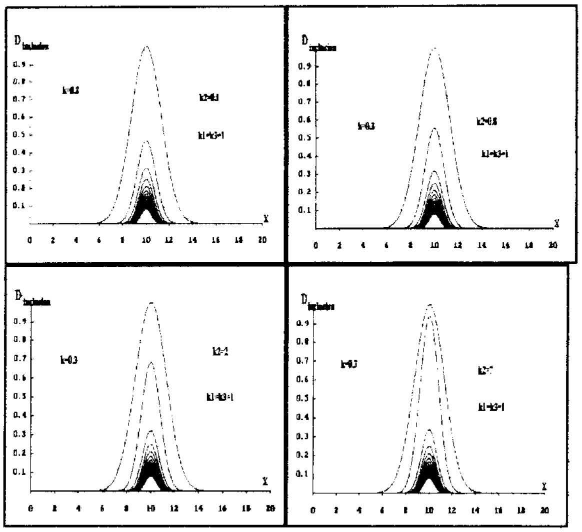

Figure 5. The influence of k2

2.2.4. *The influence of k3*

In this part, we fix k, k1, k2 and we make vary k3, in order to illustrate the influence of the function (2.4) on the degree of inclusion.

Like the previous cases, the fuzzy inclusion degree of I_1 in I_2 gets as greater as I_1 becomes closer to I_2. Its maximal value (≤ 1) is reached when the centers O1 and O2 of the two intervals coincide. It is equal to 1 when the two intervals are exactly identical, and more the difference of the sizes is important, more this value is decreases.

We remark that:

We have a very visible influence on the speed of decreasing of inclusion function. Indeed, the more k3 is small, the more the function decreases slowly. Then, once k3 becomes elevated, the function decreases quickly.

We can cancel or accented the influence of interval's position. (See figure 6)

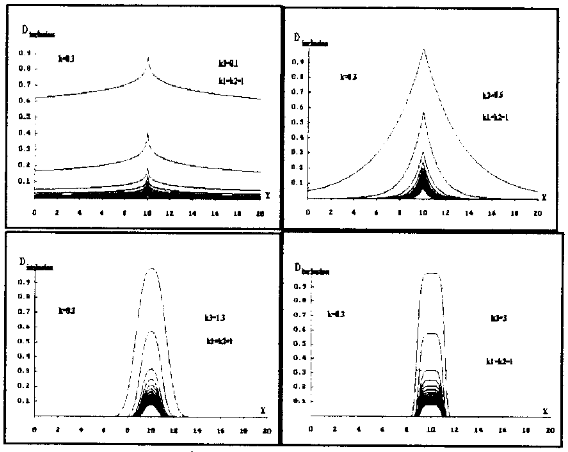

Fig. 6 The influence of k3

4. Applications

In this section, we give concrete examples to illustrate the influence of each function defining the fuzzy inclusion degree. The comparison is given on computing the inclusion degree between the first element and all elements of the considered set.

4.1. The two parity problem

The even-parity patterns have an associated output value of 1 and the odd-parity have an output of 0. We can see that a regular metric would not allow us to solve the problem of XOR.

However, if we describe this set with the interval formed between the two attributes, we can note that the intervals formed between the even parity patterns or the odd parity, have the same size. So, it's logic to cancel the influence of interval's position and accented the role of the others functions. The results of such operation are illustrated in the figures 7,8

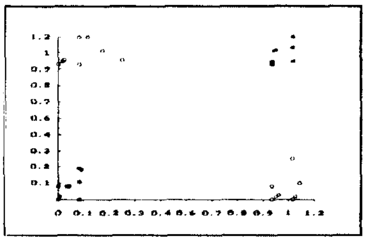

Fig. 7 A set of Xor problem

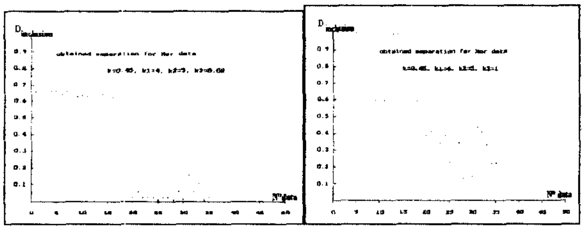

Fig. 8 Obtained separation for Xor Problem. We can see that:

- For k= 0.45, k1=4, k2=5, k3=0.02, the separation is complete, because we have relatively concealed the position of the interval.
- For k=0.45, k1=4, k2=5, k3=1, the separation becomes relatively incomplete.

4.2. The Iris data of Anderson

The Iris data of Anderson[3] has often been used as a standard set for testing the performances of data analyses algorithms and discrimination's criteria. In fact, the real structure of this set is difficult to recover. The data consists of 150 four-dimensional vectors that form three clusters. The components of a vector are the measurement of the petal length, petal width, sepal length and sepal width of a particular iris plant. We are concerned here only with the two varieties not well separated. There are 50 plants in each of the two varieties represented in the data: Virginia iris and Versicolor iris.

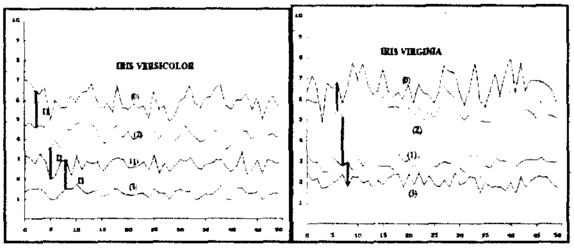

Fig. 9 The evolution of characters of two varieties of the Iris data of Anderson

By examining the graph in the figure 9, that represents the evolution of characters of the two varieties, we notice that there is a certain relationship that links parameters. Therefore, we have represented varieties by intervals Ii as it illustrated in the figure (9). We can also see that the position of the intervals can influence the result of the separation. So, we must chose adequacy the coefficients of pondering in order to translate as well the separation. The results of such operation are illustrated in the figure 10.

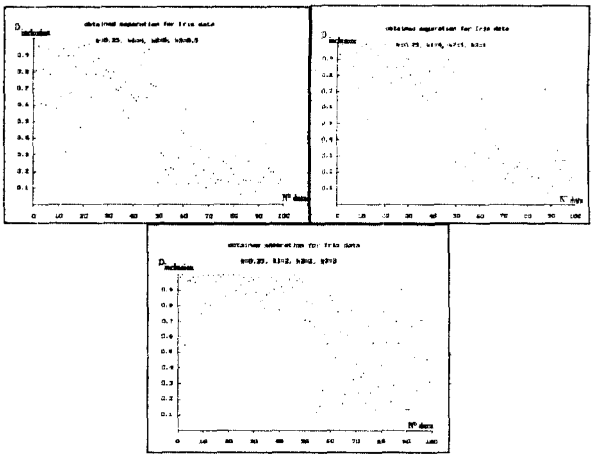

Fig. 10 Obtained separation for the Iris data. We can see that:

- For k= 0.25, k1=4, k2=5, k3=0.5, the separation is relatively complete.
- For k=0.25, k1=4, k2=5, k3=1, the separation becomes more complete.
- For k=0.25, k1=2, k2=2, k3=2, the separation is incomplete.

4.3. Construction of cluster's prototypes

Clustering has long been a popular approach to unsupervised pattern recognition. Fuzzy clustering has been shown to be advantageous over crisp one [7]. The major factor that influences the determination of appropriate groups of points is the "distance measure" chosen for the problem at hand, and the choice of representative prototype. Data is generally expressed as a collection of vectors in $\Re^n$ [8] and in most cases, the distance D is simply the Euclidean Distance in $\Re^n$.

Nevertheless, the two approaches are generally based on an adjustment of the prototype vector representing the appropriate cluster. This adjustment can falsify the representation of the class especially if the noise is present.

In this paragraph, we propose a new clustering approach in order to establish a prototype of representation that can overcome several inconvenient. This approach will include general structures. Our goal is to construct prototypes formed by fuzzy curves once the classification has advanced. So, we can supervise the results of classification by following the development of prototypes. Consequently, we can have the possibility to return back and correct if it's necessary.

4.3.1. Method principle

The principle of the method is simple. Indeed, instead of adjusting the prototype using a linear combination, we are going to link the new element to the ancient by a curve. It is similar to the process of filling a thread with adequate pearls.

This method has a lot of advantages. First, we do not lost information due to the adjustment of the prototype vector. And, we have the possibility to return back and to correct. And finally, we can supervise the functioning of the clustering because the resolution is graphic.

4.3.2. Clustering using interval's description

We use the Iris data of Anderson, because they give an illustrating example of prototype. We propose to represent the varieties using the intervals [3] as it illustrated in the figure 9. We have not added any ad hoc information. So, in stead of representing the elements with a vector formed by the four characters (c1, c2, c3, c4), we will represent it with a vector formed by six intervals (I_1, I_2,..., I_3) who represents the width of the intervals formed by the superposition of the four characters[2]. And for the discrimination, we compare the degree of fuzzy inclusion between the defined intervals.

4.3.3. Construction of prototypes in sequential manner

The construction of prototypes is made element by element and in sequential manner. The general form of the algorithm is like following:

Fix the threshold, initialize the number of cluster C=1;

Initialize the prototype of the first cluster with the first element

for i=2 to number of elements
for j=1 to C
compute the fuzzy inclusion degree
seek the maximum degree with its prototype
if (maximum > threshold)
link the element to this prototype by a curve
else
initialize another prototype
c=c+1.

The obtained results are illustrated in the following figures. We can see that the number of prototypes depends on threshold. In addition, the last part of the first prototype (the framed one) is not conform with the shape of all the fuzzy curve. So, we must search another algorithm that can give us a unique prototype.

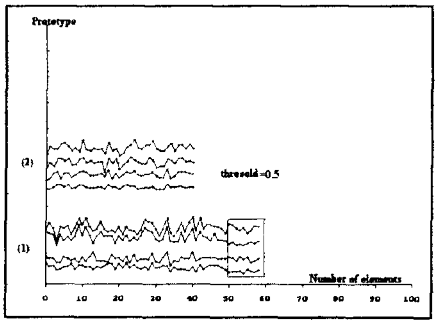

Fig. 11 Obtained prototype for the Iris data with threshold =0.5

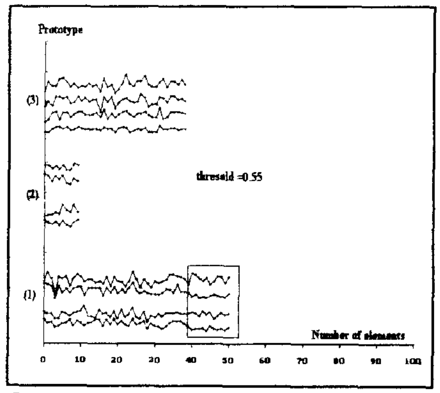

Fig. 12 Obtained prototype for the Iris data with threshold =0.55

Remark

We put in order all elements of the two clusters before using them in the algorithm. So, we can say that this method depends on the order of apparition of the elements. However, even if the elements are put in order we have not succeed to separate them.

4.3.4. Construction of prototypes in selective manner

The construction of prototypes is made by choosing the most similar elements. The general form of the algorithm is like following.

First construction

Fix a threshold very high (>0.7), initialize the number of cluster C=1;
Initialize the prototype of the first cluster with the two most similar elements
repeat
seek the closer elements to the prototype C
if (inclusion degree > threshold)
link the element to this prototype by a curve
else
initialize another prototype with the two similar elements (in the set of elements not used)
C=C+1.
Until (number of elements)
second construction
regroup the most similar prototype in one curve
or regroup the similar prototype in one cluster

The result of the first construction is as follows.

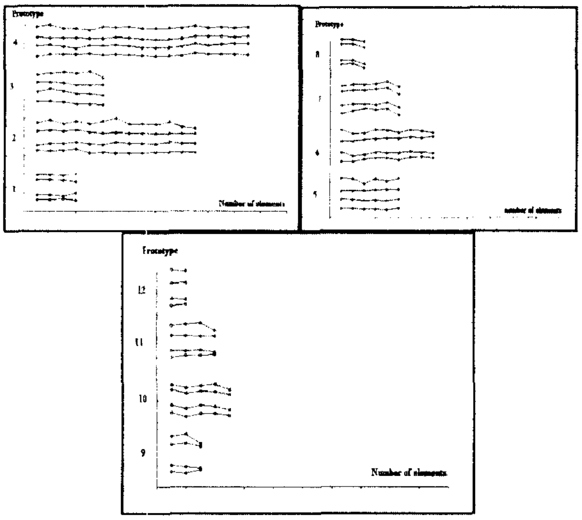

Fig. 13 Obtained prototypes for the Iris data

This first construction tries to gather the most similar elements in one prototype. So, different prototypes can represent the same cluster.(See prototypes in following figures (2,3,4,5)). This construction is unique.

The second construction is illustrated in the following figure.

We undertake this construction using two manners. Indeed, we can try to gather the most similar prototypes in one prototype or we can let the prototypes without linking them and allocate them in one cluster.

In this case, we should define an inclusion degree relative to each element in each cluster. This inclusion degree is the maximum value that it can take in each prototype that represent the cluster. Thus, if we want to verify the affiliation of an element in a cluster we should seek the inclusion degree of this element in cluster: the element will be member of the cluster giving the higher inclusion degree.

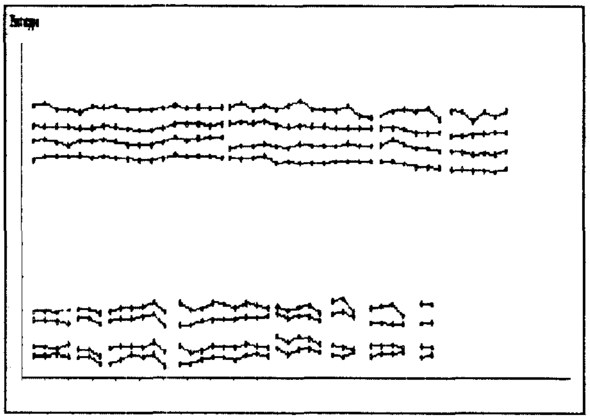

Fig. 14 Obtained final prototype for the Iris data

5. References and bibliography

1. F. KARBOU, Nouveaux Concepts Flous en Analyse de Données pour le Codage, la Similarité et la Représentation Graphique des variétés agricoles, CIMASI'96, p.155-163, ENSEM, CASABLANCA, Morocco, 14-16 Novembre (1996)
2. F. KARBOU, New Fuzzy Concepts, FUZZ-IEEE'97, Six IEEE international Conference on Fuzzy Systems, Barcelona, Spain, 1-5 JULLY, 1997.
3. Michael P. WINDHAM, Cluster Validity for Fuzzy Clustering Algorithm, Fuzzy sets and systems 5(1981), 177-185, North-Holland publishing Company.