NEW RESULTS IN FUZZY CLUSTERING BASED ON THE CONCEPT OF INDISTINGUISHABILITY RELATION

R. López de Mántaras

Facultat d'Informàtica
Universitat Politècnica de Barcelona
Dulcet, 12. Barcelona-34. Spain.

*L. Valverde**

Dept. de Matemàtiques i Estadística
Universitat Politècnica de Barcelona
Avda. Diagonal, 649. Barcelona-28. Spain.
Barcelona. Spain

KEY WORDS

Fuzzy Clustering, Cluster Validity, Indistinguishability Relation, t-norm, t-conorm, G-pseudometric.

1. INTRODUCTION

One of the fundamental problems in clustering is its validation. This issue has been addressed by many researchers in different ways depending on the methods they use to solve the clustering problem, nevertheless the existing approaches are based on the introduction of a validity function which is intented to measure the validity of the clustering obtained by the different methods.

In the case of hierarchical methods, the validity function measures the compacity of the obtained partition. The methods based on graph theory lead for example to measures of connectivity, length of the chains, etc.. The methods based on objective functions usually use the objective function itself as a validity function. Other well-known methods in the setting of fuzzy clustering are, the "degree of separability", the "partition coefficient", the "classification entropy" (Bezdek,1981). Such a variety of method-dependent and "ad hoc" measures of validity suggests that it is very difficult to obtain more general solutions to the "validity problems". But, before addressing the validity aspect we think that it is necessary to define **what is a cluster?** and only after that it makes sense to talk about **how good is a cluster?**.

In our work we first address this issue and we give a new definition of fuzzy r-cluster that extend Ruspini's definition (Ruspini, 1982). Our definition is based on the new concept of indistinguishability relation (Trillas, 1982) which includes, as particular cases the concepts of similarity relation of Zadeh (Zadeh, 1971), probabilistic relation of Menger (Menger, 1951) and likeness relation of Ruspini (Ruspini, 1982).

(*) Visiting Scholar at the Computer Science Division of the University of California at Berkeley. Research partially supported by a fellowship of the **Comissió Interdepartamental per a la Recerca i Innovació Tecnològica** (CIRIT) of the Generalitat de Catalunya, and by NSF Grant ECS-8209679 and NESC Contract N0039-83-C-0243.

The main difficulty with the clustering process is that the data and the obtained partition are "separated" by the algorithm that built such partition, consequently it is generally imposible to gain some insight about the "structure" of the data based only on the information available observing the obtained partition; and the existing methods do no look in general to the data but only the partition.

The approach presented here is heavily based on the data themselves. First we construct an indistinguishability relation among a data set (that includes the prototypes of the clusters) based on the degree of membership, to the different clusters, of each element of the data set. Finally, we measure the validity comparing (through a distance measure) the degrees of membership of the data to the clusters, with the degrees of indistinguishability between the data and the prototypes of each cluster. The basic idea is that *the smaller the difference between the degrees of membership and the degrees of indistinguishability, the better the clustering.* The intuitive reason can be expressed as follows: The degrees of membership of the elements to a cluster depend only on that cluster but, the degrees of indistinguishability between the elements and the prototypes of the clusters depend on the values of the degrees of membership of these elements and these prototypes with respect *to all the clusters*. Therefore. if these degrees are the same, the membership functions constitute a fuzzy cluster coverage (the counterpart, for indistinguishability relations, of partitions). And according to our definition, the clustering obtained is valid if the membership functions form a fuzzy cluster coverage. We have implemented an algorithm that, given the membership functions, measures the clustering validity and we give some results obtained with some examples.

We also suggest that the validity measure obtained can be used to reconsider initial decisions about the election of the prototypes, the number of clusters, etc.. Also since, as we shall see, and as a "side effect" we get information about the logical and metric properties of the data, we could use such information in order to make a geometrical representation of the data (factor analysis, principal components, etc.). We start giving some definitions and theorems whose proofs are omitted because of space reasons (See Valverde, 1983).

2. ON INDISTINGUISHABILITY RELATIONS

2.1. A general definition of indistinguishability relations and their use in different fields can be found in (Trillas, 1982; Trillas and Valverde, 1983). Here we will be mainly concerned with a special kind of these relations: the F-indistinguishability relations.

Definition 1. Given a non-empty set X, a map E from XxX into [0,1] is called **F-indistinguishability relation** if the following properties hold for any x,y and z in X:

i) $E(x,y)=1$,

ii) $E(x,y)=E(y,x)$, and

iii) $F(E(x,y),E(y,z)) \leq E(x,z)$.

That is, F-indistinguishability relations are simply fuzzy binary relations which are reflexive, symmetric and F-transitive, F being a continuous t-norm. Thus, similarity relations (Zadeh 1971) are F-indistinguishability relations with $F(x,y)=\mathrm{Min}(x,y)$; the same applies to probabilistic relations of indistinguishability (Menger 1951) with $F(x,y)=x.y$ and likenes relations (Ruspini 1982) with $F(x,y)=\mathrm{Max}(x+y-1,0)$.

It is well known that likeness relations were introduced in the framework of fuzzy clustering and, among other considerations, the rationale behind their definition is given by the fact: if E is a likeness relation on X, then $d(x,y)=1-E(x,y)$ is a normalized pseudo-metric on X; that is such a relation is given by (and gives a) distance between the elements of X. In this way the notion of distance, which has long been used in many contexts as a measure of similarity, falls into the scope of some F-indistinguishability relations. Moreover, we have

Theorem 1. Let d be a pseudo-distance on a set X, then

a) $L_d(x,y) = 1/(1+d(x,y))$ is a likeness relation on X, and viceversa.

b) $P_d(x,y) = \exp(-d(x,y))$ is a probabilistic relation on X, and vice versa.

F-indistinguishability relations with respect to t-norms different from the the t-norms corresponding to likeness and probabilistic relations do not, in general, give a pseudo-metric but a G-pseudometric:

Definition 2. Given a continuous t-conorm G, a map from XxX into [0,1] is called a **G-pseudometric** if the following properties hold for any x,y and z in X:

i) m(x,x)=0,

ii) m(x,y)=m(y,x), and

iii) G(m(x,y),m(y,z)) $\geq$ m(x,z).

It can easily be seen that, for any continuous t-norm G,

is a G-pseudometric on [0,1], where f is a continuous and strictly decreasing bijection from [0,1] into itself, and $G^{\wedge}$ is the quasi-inverse of G, i.e. $G^{\wedge}(x|y) = \mathrm{Inf}\{a \in [0,1] \mid G(a,x) \geq y\}$

The links between G-pseudometrics and F-indistinguishability relations are described in the following theorem:

Theorem 2. Let m be a G-pseudometric on a set X. Then, for any continuous and strictly decreasing bijection, f, from [0,1] into itself, $E(x,y) = f(m(x,y))$ is an F-indistinguishability relation with respect to $F(x,y) = f^{-1}(G(f(x),f(y))$ and vice versa.

That is, any F-indistinguishability relation on a set X is determined, up to an order-reversing bijection on the unit interval, by a G-pseudometric and vice versa.

2.2. In (Ruspini 1982) the concept of fuzzy r-cluster is introduced in the following way: given a reflexive and symetric fuzzy relation r on X, a fuzzy subset g of X is called *fuzzy r-cluster* if the following properties hold for any x,y and z in X:

a) If g(x)=1 then g(y)=r(x,y),

b) $|g(x) - g(y)| \leq 1 - r(x,y).$

It is shown that the family of similarity classes given by any likeness relation r (i.e. the fuzzy subsets of X defined by $g_x(y) = r(x,y)$) satisfies the two properties:

i) $\sup_{x \in X} g_x(y) = 1$ and $\sup_{y \in X} g_x(y) = 1$ (i.e. $\{g_x\}_{x \in X}$ is a *fuzzy coverage* of X).

ii) For any $x \in X$, g_x is a fuzzy r-cluster.

Conversely, if for a given reflexive and symmetric relation r there exists a fuzzy coverage of X satisfying the above property, then r should be a likeness relation.

Thus, fuzzy r-clusters are the counterpart, for likeness relations, of classical clusters (equivalence classes with respect to an equivalence relation). In order to extend these results to any F-indistinguishability relation, let be noticed that Ruspini's definition involves a particular metric in the unit interval, the restriction to the Euclidean distance. In fact, this definition may be viewed as a generalization of the definition of classical clusters because, at is easy to show, classical clusters can be characterized by means of

a) If $\mu_A(x) = 1$ then $\mu_A(y) = \mu_R(x,y)$,

b) $d_o(\mu_A(x), \mu_A(y)) \leq 1 - \mu_R(x,y)$,

where μ_A is the characteristic function of the set A and d_o is the discrete distance on the two-point set {0,1}.

Thus, by taking different metrics, different kinds of fuzzy r-clusters and, therefore, different kinds of F-indistinguishability relations r will be obtained:

Definition 3. Let r be a reflexive and symmetric fuzzy relation on X. For a given continuous t-conorm G, a fuzzy subset g of X will be termed **fuzzy r-cluster with respect to** $m_G(a,b) = G^{\wedge}(f(\mathrm{Max}(a,b)) \mid f(\mathrm{Min}(a,b)))$, if the following properties hold for any x and y in X:

a) If g(x)=1 then g(y)=r(x,y),

b) $m_G(g(x),g(y)) \le f(r(x,y))$.

A **fuzzy r-cluster coverage** will be a fuzzy coverage of X such that each of its elements is a fuzzy r-cluster.

This definition extends Ruspini's definition, that is we have:

Theorem 3. For a given F-indistinguishability relation r on X, the family $C_r = \{g_x \in [0,1]^X \mid x \in X\}$, where $g_x(y) = r(x,y)$, is a fuzzy r-cluster coverage of X with respect to m_G, where $G(x,y) = f^{-1}(F(f(x),f(y)))$.

Conversely, we have

Theorem 4. Let C be a fuzzy coverage of X and r a reflexive and symmetric fuzzy relation on X. If every element of C is a fuzzy r-cluster with respect to m_G, then r is a F-indistinguishability relation, where $F(x,y) = f(G(f^{-1}(x),f^{-1}(y)))$.

EXAMPLES

(E1) If $G(a,b) = \mathrm{Min}(a+b,1)$ and $f(a) = 1-a$, then

and r is a likeness relation.

(E2) If $G(a,b) = \mathrm{Max}(a,b)$, then

and r is a similarity relation.

(E3) If $G(a,b) = a+b-ab$ and $f(a) = 1-a$, then

and r is a probabilistic relation.

2.3. Our approach is heavily based on the representation theorems for F-indistinguishability operators (Ovchinnikov, 1982; Valverde, 1982) which provide a way to construct F-indistinguishability operators. The following is the second author formulation of that result:

Theorem 5.(Representation theorem). Let r be a reflexive fuzzy relation on X and let F be a continuous t-norm. Then r is a F-indistinguishability relation on X if, and only if, there exists a family of fuzzy subsets of X, $(h_j)_{j \in J}$, such that

where $F^{\wedge}$ is the quasi-inverse of F, i.e. $F^{\wedge}(x \mid y) = \mathrm{Sup}\{a \in [0,1] \mid F(a,x) \le y\}$.

Consequently, given a continuous t-norm F, we can associate a F-indistinguishability relation with any family of fuzzy subsets of X and, therefore, a fuzzy r-cluster coverage with respect to some G-metric of X is obtained.

EXAMPLES

(E4) If $F(a,b) = \mathrm{Min}(a,b)$, then the F-indistinguishability relation associated to $(h_j)_{j \in J}$, $h_j \in [0,1]^X$, is given by

where $J_{xy} = \{ j \in J \mid h_j(x) \neq h_j(y) \}$.

(E5) If $F(a,b) = a.b$, then

(E6) If $F(a,b) = \mathrm{Max}(a+b-1,0)$, then

We remark that, since F-indistinguishability relations and G-pseudometrics are dual concepts, theorem 5 provides also a representation for any G-pseudometric.

3. CLUSTER VALIDITY

The previous theorem can be rephrased as follows: To any family of fuzzy subsets (membership functions) of X and for any continuous t-norm F, we can associate a **fuzzy cluster coverage** of X and, by duality, a G-pseudometric, i.e. a distance between the elements of X.

The basic principle underlying the cluster validity is given in the following

Proposition. If the membership functions $\{h_j\}_{j \in J}$ constitute a fuzzy cluster coverage, then there exists a t-norm, F, such that the degree of membership $h_j(x)$ of an element x with respect to the cluster j is the same as the degree of indistinguishability (with respect to the indistinguishability operator generated by F) between this element x and the prototype of the cluster j.

In general the family $\{h_j\}_{j \in J}$ of membership functions do not constitute a fuzzy coverage and, therefore, the degrees of membership and the degrees of indistinguishability will not coincide. Then we state the following

Principle The best t-norm to be used to generate the fuzzy coverage associated with the family $\{h_j\}_{j \in J}$ will be that which corresponds to *the smallest distance* (using the corresponding m_G) between the $\{h_j\}_{j \in J}$ and the correspondig $\{g_j\}_{j \in J}$ in the indistinguishability relation.

Remark. We detect at the same time the best *logic* (best t-norm, best t-conorm) and the best *topology* (m_G) underlying the structure of the data set X.

ALGORITHM. Given a family $\{F_\lambda\}_{\lambda \in \Gamma}$ of t-norms such that $F_\lambda \geq F_{\lambda'} \text{ if } \lambda \leq \lambda'$, and the family $\{h_j\}_{j \in J}$ of membership functions, do the following for each t-norm:

(1) Construct the corresponding F_λ-indistinguishability relation.

(2) For each cluster j, calculate the local distance

where |X| is the cardinal of the data set X and $ m sub {G sub lambda} $ is the distance associated with the t-norm $ F sub lambda $ (for a fixed order-reversing bijection f form the unit interval into itself).

(3) Calculate the global distance

where |J| is the number of classes.

(4) Select the set of $ lambda 's $ such that $ D sup lambda $ is minimum, and among these, the biggest one. The reason for choosing the biggest $ lambda $ is the following: if $ lambda <= lambda prime $ then $ F sub lambda >= F sub lambda prime $, therefore $ G sub lambda <= G sub lambda prime $ and this implies that $ m sub {G sub lambda} >= m sub {G sub {lambda prime}} $ (as it has been shown in (Valverde 1983)). That is, a fuzzy coverage with respect to $ F sub lambda $ is also a fuzzy coverage with respect to $ F sub {lambda prime} $, but the values of the distances measured using $ m sub {G sub lambda} $ are bigger, in other words we say that$ ~~m sub {G sub lambda} $ has a higher resolution power than $ m sub {G sub {lambda prime}} $.

EXAMPLES

We have tried our appoach with several examples involving two clusters, and we have considered 3 different t-norms:

$ 1)~F(x,y)^=^Min(x,y) $
$ 2)~F(x,y)^=^x.y $
$ 3)~F(x,y)^=^Max(x+y-1,0) $.

The order-reversing bijection from the unit interval into itself we have used is $ f(x)^=^1-x $, so the corresponding t-conorms are:

$ 1 prime)~G(x,y)^=^Max(x,y) $
$ 2 prime)~G(x,y)^=^x+y-x.y $
$ 3 prime)~G(x,y)^=^Min(x+y,1) $.

In almost all cases, the best t-norm is the last one of the previous three, and corresponds precisely to the logic of Lukasiewicz Aleph-1, and the corresponding $ m sub G $ is the restriction to the unit interval of the Euclidean distance.

allbox; c s s c c c c n n. DATA SAMPLE #1 S $ h sub 1 $$ h sub 2 $ $ x sub 1 $0.10 0.60 $ x sub 2 $0.60 0.10 $ x sub 3 $0.10 0.90 $ x sub 4 $0.30 0.40 $ x sub 5 $0.30 0.90 $ x sub 6 $0.10 0.60 $ x sub 7 $0.60 0.10 $ x sub 8 $0.40 0.20 $ x sub 9 $0.30 0.80

```
allbox; c c c c s s s s s s c c c c c c c c c c n n n n n n n n n. DS #1    $ g1l $    $ g2l
$    $ F(x,y)^=^Min(x,y) $ r($ xi $,$ xj $)   $ x1 $    $ x2 $    $ x3 $    $ x4 $    $ x5
$    $ x6 $    $ x7 $$ x8 $$ x9 $ $ x1 $    1   0.1   0.1   0.1   0.1   1    0.1   0.1   0.1
$       x2        $      0.1   1    0.1   0.1   0.1   0.1   0.1   0.1   0.1          $        x3
$    0.1   0.1   1     0.1   0.1   0.1   0.1   0.1   0.1                 $                  x4
$    0.1   0.1   0.1   1     0.4   0.1   0.1   0.2   0.4                 $                  x5
$    0.1   0.1   0.1   0.4   1     0.1   0.1   0.2   0.8                 $                  x6
$    1     0.1   0.1   0.1   0.1   1     0.1   0.1   0.1                 $                  x7
$    0.1   0.1   0.1   0.1   0.1   0.1   1     0.1   0.1                 $                  x8
$    0.1   0.1   0.1   0.2   0.2   0.1   0.1   1     0.2                 $                  x9
$    0.1   0.1   0.1   0.4   0.8   0.1   0.1   0.2   1
```

```
Cluster #1: $ d sub 1 sup lambda ^=^0.90 $
Cluster #2: $ d sub 2 sup lambda ^=^0.90 $
$ D sup lambda ^=^0.90 $
```

```
allbox; c c c c s s s s s s c c c c c c c c c c c c n n n n n n n n n. DS #1      $ gl1 $      $ g2l
$      $ F(x,y)^=^xy  $ r($ xi $,$ xj $)  $ x1 $       $ x2 $       $ x3 $       $ x4 $       $ x5
$      $ x6 $       $ x7 $$ x8 $$ x9 $ $ x1 $       1     0.17  0.67  0.33  0.33  1     0.17  0.25  0.33
$         x2        $      0.17  1     0.11  0.25  0.11p0.17  1     0.5   0.13        $         x3
$      0.67  0.11  1     0.33  0.33  0.67  0.11  0.22  0.33                 $                  x4
$      0.33  0.25  0.33  1     0.44  0.33  0.25  0.5   0.5                  $                  x5
$      0.33  0.11  0.33  0.44  1     0.33  0.11  0.25  0.89                 $                  x6
$      1     0.17  0.67  0.33  0.33  1     0.17  0.25  0.33                 $                  x7
$      0.17  1     0.11  0.25  0.11  0.17  1     0.5   0.13                 $                  x8
$      0.25  0.5   0.22  0.5   0.25  0.25  0.5   1     0.25                 $                  x9
$      0.33  0.13  0.33  0.5   0.89  0.33  0.13  0.25  1
```

```
Cluster #1: $ d sub 1 sup lambda ^=^0.446 $
Cluster #2: $ d sub 2 sup lambda ^=^0.685 $
$ D sup lambda ^=^0.565 $
```

```
allbox; c c c c s s s s s s c c c c c c c c c c c c n n n n n n n n n. DS #1      $ gl1 $      $ g2l
$      $ F(x,y)^=^Max(x+y-1,0)  $ r($ xi $,$ xj $)  $ x1 $       $ x2 $       $ x3 $       $ x4
$      $      x5      $      $      x6      $      $      x7      $$      x8      $$      x9      $      $      x1
$      1     0.5   0.7   0.8   0.7   1     0.5   0.6   0.8                  $                  x2
$      0.5   1     0.2   0.7   0.2   0.5   1     0.8   0.3                  $                  x3
$      0.7   0.2   1     0.5   0.8   0.7   0.2   0.3   0.8                  $                  x4
$      0.8   0.7   0.5   1     0.5   0.8   0.7   0.8   0.6                  $                  x5
$      0.7   0.2   0.8   0.5   1     0.7   0.2   0.3   0.9                  $                  x6
$      1     0.5   0.7   0.8   0.7   1     0.5   0.6   0.8                  $                  x7
$      0.5   1     0.2   0.7   0.2   0.5   1     0.8   0.3                  $                  x8
$      0.6   0.8   0.3   0.8   0.3   0.6   0.8   1     0.4                  $                  x9
$      0.8   0.3   0.8   0.6   0.9   0.8   0.3   0.4   1
```

```
Cluster #1: $ d sub 1 sup lambda ^=^0.467 $
Cluster #2: $ d sub 2 sup lambda ^=^0.533 $
$ D sup lambda ^=^0.500 $
```

allbox; c s s c c c c c n n. DATA SAMPLE #2 S $ h sub 1 $$ h sub 2 $ $ x1 $ 1 0 $
x2 $ 0 1 $ x3 $ 0.80 0.10 $ x4 $ 0.90 0.20 $ x5 $ 0.50 0.90 $ x6
$ 0.40 0.90 $ x7 $ 0.10 0.80 $ x8 $ 0.50 0.70

allbox; c c c c s s s s s c c c c c c c c c c c n n n n n n n n. DS #2 $ g11 $ $ g21 $ $
F(x,y)ˆ=ˆMin(x,y) $ r($ xi $,$ xj $) $ x1 $ $ x2 $ $ x3 $ $ x4 $ $ x5 $ $
x6 $ $ x7 $$ x8 $ $ x1 $ 1 0 0 0 0 0 0 0 $ x2
$ 0 1 0 0 0 0 0 0 $ x3
$ 0 0 1 0.1 0.1 0.1 0.1 0.1 $ x4
$ 0 0 0.1 1 0.2 0.2 0.1 0.2 $ x5
$ 0 0 0.1 0.2 1 0.4 0.1 0.7 $ x6
$ 0 0 0.1 0.2 0.4 1 0.1 0.4 $ x7
$ 0 0 0.1 0.1 0.1 0.1 1 0.1 $ x8
$ 0 0 0.1 0.2 0.7 0.4 0.1 1

Cluster #1: $ d sub 1 sup lambda ˆ=ˆ0.75 $

Cluster #2: $ d sub 2 sup lambda ˆ=ˆ0.75 $

$ D sup lambda ˆ=ˆ0.75 $

allbox; c c c c s s s s s c c c c c c c c c c c n n n n n n n n. DS #2 $ g11 $ $ g21 $ $
F(x,y)ˆ=ˆxy $ r($ xi $,$ xj $) $ x1 $ $ x2 $ $ x3 $ $ x4 $ $ x5 $ $ x6
$ $ x7 $$ x8 $ $ x1 $ 1 0 0 0 0 0 0 0 $ x2
$ 0 1 0 0 0 0 0 0 $ x3
$ 0 0 1 0.5 0.11 0.11 0.13 0.14 $ x4
$ 0 0 0.5 1 0.22 0.22 0.13 0.29 $ x5
$ 0 0 0.11 0.22 1 0.8 0.2 0.78 $ x6
$ 0 0 0.11 0.22 0.8 1 0.25 0.78 $ x7
$ 0 0 0.13 0.13 0.2 0.25 1 0.2 $ x8
$ 0 0 0.14 0.29 0.78 0.78 0.2 1

Cluster #1: $ d sub 1 sup lambda ˆ=ˆ0.40 $

Cluster #2: $ d sub 2 sup lambda ˆ=ˆ0.45 $

$ D sup lambda ˆ=ˆ0.425 $

```
allbox; c c c c s s s s s s c c c c c c c c c c c n n n n n n n n. DS #2   $ g11 $   $ g21 $   $
F(x,y)^=^Max(x+y-1,0) $ r($ xi $,$ xj $)  $ x1 $   $ x2 $   $ x3 $   $ x4 $   $ x5
$   $ x6 $   $ x7 $$ x8 $ $ x1 $   1   0   0.8   0.8   0.1   0.1   0.1   0.3 $ x2
$   0     1     0.1   0.1   0.5   0.6   0.8   0.5           $   x3
$   0.8   0.1   1     0.9   0.2   0.2   0.3   0.4           $   x4
$   0.8   0.1   0.9   1     0.3   0.3   0.2   0.5           $   x5
$   0.1   0.5   0.2   0.3   1     0.9   0.6   0.8           $   x6
$   0.1   0.6   0.2   0.3   0.9   1     0.7   0.8           $   x7
$   0.1   0.8   0.3   0.2   0.6   0.7   1     0.6           $   x8
$   0.3   0.5   0.4   0.5   0.8   0.8   0.6   1
```

```
Cluster #1: $ d sub 1 sup lambda ^=^0.125 $
Cluster #2: $ d sub 2 sup lambda ^=^0.125 $
$ D sup lambda ^=^0.125 $
```

Concluding remarks.

We have generalized some previous results of Ruspini and others in fuzzy clustering, using the new concept of indistinguishability relation based on the concept of t-norm and also we have studied its metrical properties through the dual concept of t-conorm that leads to G-pseudometrics. From the concept of G-pseudometric we have defined fuzzy r-clusters and fuzzy cluster coverages. Finally, we have proposed a measure of cluster validity based on the concept of fuzzy coverage.

It is important to notice that the process of measuring the validity is carried out before any decision concerning the assignments of elements to the clusters. Therefore we can postpone the decision step until we have an acceptable cluster validity. That is, we have a sort of closed loop in the sense that a bad measure of validity obliges the user to reconsider previous hypothesis like, for example, the number of clusters, the values of the prototypes, etc.. Right now we are studying this "close loop" aspect in the setting of different classification algorithms.

REFERENCES

Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum Pub. Co., 272 pages.

Bezdek, J.C. and Harris, J.O. (1978). Fuzzy Partitions and Relations: An Axiomatic Basis for Clustering. *Fuzzy Sets and Systems 1, 112-127.*

Menger, K. (1951). Probabilistic Theory of Relations. *Proc. Nat. Acad. Sci. USA, 37, 178-180.*

Ovchinnikov, S.V. (1982). On Fuzzy Relational Systems. *Proc. 2nd. World Conference on Mathematics at Service of Man.* Univ. Politécnica Las Palmas de Gran Canaria. Spain. 566-568.

Ruspini, E. (1982). Recent Developments in Fuzzy Clustering. In R.R. Yager (Ed.): *Fuzzy Set and Possibility Theory: Recent Developments.* Pergamon Press, 133-147.

Trillas, E. (1982). Assaig sobre les relacions d'indistingibilitat. *Actes del Primer Congrès Català de Lògica Matemàtica,* Barcelona, 51-59.

Trillas, E. and Valverde, L. (1983). An Inquiry on Indistinguishability Operators. To appear in H.J. Skala et al. (Eds.): *Aspects of Vagueness.* Reidel.

Valverde, L. (1982). *Contribució a l'estudi dels models matemàtics per lògiques multivalents.* Ph.D. Thesis. Universitat Politècnica de Barcelona.

Valverde, L. (1983). On the Structure of F-Indistinguishability Operators. Mimeo.

Zadeh, L.A. (1971). Similarity Relations and Fuzzy Orderings. *Inf. Sci. 3, 177-200.*