Discrimination power of measures of comparison

Maria Rifqi, Vincent Berger, Bernadette Bouchon-Meunier

Discrimination power of measures of comparison

M. Rifqi [a], V. Berger [b] and B. Bouchon-Meunier [a]

[a] *LIP6 – Université Pierre et Marie Curie*
Case 169 – 4, Place Jussieu, 75252 Paris Cedex 05
{rifqi, bouchon}@poleia.lip6.fr

[b] *Laboratoire Central de Recherche – Thomson-CSF*
Domaine de Corbeville, 91400 Orsay
berger@thomson-lcr.fr

Abstract

This paper is based on a framework [3] for a formalization of measures of comparison of fuzzy objects. The purpose is to describe the behaviour of measures of comparison within a given family in order to facilitate the choice of a particular measure. It can be done thanks to the discrimination power of a measure.

1 Introduction

The comparison of two objects is a usual task for many and various domains as psychology, statistics, fuzzy sets theory, ... Indeed, comparisons are useful in classification for the matching step, in clustering for the construction of classes [6], [7], in decision-making for the search of the best candidate, ...

Comparisons are usually realized by means of a measure of comparison. The used measure is often a distance. But, more and more, a similarity or a dissimilarity measure is chosen. But the choice of an appropriate measure among all avalaible measures in literature is not an easy task. It is linked to the problem of the characterization of relevant properties for the considered task. We have proposed a framework in order to deal with measures of comparison [3]. This framework displays the main families of measures of comparison according to the properties they satisfy. Hence, the existing measures of comparison can be classified. And it is known that a classification simplify a problem.

However, the problem of the choice of a measure of comparison *within* a same family is still present. This paper proposes a solution to this problem.

The proposed solution lies in the discrimination power of a measure. This way, measures can be compared among themselves according to their behaviour. The analysis of behaviours of measures of comparison is easy thanks to a geometrical interpretation. This geometrical interpretation is obvious if variables stepping in family of measures are normalized.

2 Measures of comparison

In [3], we have proposed to formalize a measure of comparison between two fuzzy sets as a function of the common elements and the distinctive elements.

Formally, for any set Ω of elements, let $F(\Omega)$ denote the set of fuzzy subsets of Ω.

Definition 1 *A* fuzzy set measure M *is supposed to be given, that is to say a mapping defined on* $F(\Omega)$ *and taking values in* $\mathbb{R}^+$ *such that, for all* A *and for all* B *in* $F(\Omega)$*:*

MI1 *:* $M(A) = 0 \Leftrightarrow A = \emptyset$.

MI2 *: if* $B \subseteq A$*, then* $M(B) \leq M(A)$.

A fuzzy set measure is close to the definition of an existential evaluator given by [9].

Definition 2 *An* M-measure of comparison *on* Ω *is a mapping* $S : F(\Omega) \times F(\Omega) \rightarrow [0,1]$ *such that* $S(A,B) = F_S(M(A \cap B), M(B - A), M(A - B))$*, for a given mapping* $F_S : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow [0,1]$ *and a fuzzy set measure* M *on* $F(\Omega)$.

We denote:

- $X = M(A \cap B)$
- $Y = M(B - A)$
- $Z = M(A - B)$

We are interested in measures of comparison which evaluate the likeliness of two descriptions. We have called them *measures of similitude.*

Definition 3 *An* M-measure of similitude S *on* Ω *is an* M*-measure of comparison* S *such that* $F_S(X, Y, Z)$ *is non decreasing with respect to* X*, non increasing with respect to* Y *and* Z*.*

Remark Tversky's contrast model [14] is compatible with M-measures of similitude. Indeed, Tversky has given a relation which can be generalized to fuzzy sets [12], [3] as follows:

$$S(A, B) = f(A \cap B)/(f(A \cap B) + \alpha f(A - B) + \beta f(B - A))$$

with $\alpha, \beta \geq 0$. This quantity is an f-measure of similitude if f is a fuzzy set measure.

M-measures of similitude can be distinguished more subtly in three types: measures of satisfiability, measures of resemblance and measures of inclusion. In this paper, we focus on the two first types of measures of similitude.

2.1 Measures of satisfiability

A measure of *satisfiability* corresponds to a situation in which we consider a reference object or a class and we need to decide if a new object is compatible with it or satisfies it. More particularly, measures of satisfiability are appropriate for rule base systems. For example, in [2] or in [1] objects are classified by means of a decision tree. In a decision tree, a node represents a test on the chosen attribute during the learning stage; each edge of this node is associated with a value of the attribute. The classification of a new object comes to find consecutive edges from the root to the leaves. In [2] and in [1], the comparison between the value of an attribute of the new example with test-values associated with each edge is realized by means of a measure of satisfiability.

The satisfiability of a reference description A of $F(\Omega)$ by a new description B defined as a fuzzy subset of Ω has been defined as follows:

Definition 4 *An* M-measure of satisfiability *on* Ω *is a measure* $S(A, B) = F_S(X, Y, Z)$ *such that:*

- F_S *is increasing with respect to* X*, decreasing with respect to* Y *and independant of* Z

Let us denote $F_{S\cdot}(X, Y) = F_S(X, Y, Z)$*. Then,*

- $F_{S\cdot}(X, 0) = 1$ *for all* X

- $F_{S.}(0, Y) = 0$ *for all* $Y \neq 0$

With this definition, satisfiability can *a priori* be different with two pairs of fuzzy sets distinctive only because of the scale. It is desirable that a satisfiability measure depends only on the relative weights of X and Y and not on the scale of the system. In order to obtain an objective measure, we propose to normalize the satisfiability measure.

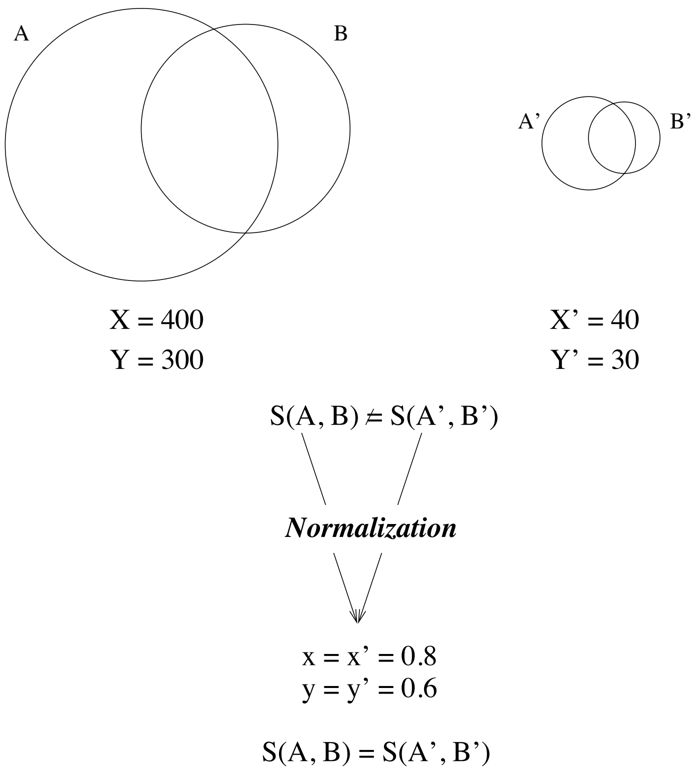

Figure 1: The effect of the normalization on measures of satisfiability

We consider:

$$\begin{aligned} x &= \frac{X}{\sqrt{X^2+Y^2}} \quad \text{the reduced intersection} \\ y &= \frac{Y}{\sqrt{X^2+Y^2}} \quad \text{the reduced distinctive feature} \end{aligned}$$

Figure 1 shows the differences of behaviour between the two definitions of measures of satisfiability.

As $x^2 + y^2 = 1$, the domain of definition of the measure of satisfiability is a quarter of circle. It can be described by a unique argument ϕ, with $\phi = \arctan \frac{y}{x}$. We denote the measure of satisfiability $S(A, B) = \eta(\phi)$.

The conditions of definition 4 become:

- η is decreasing with respect to ϕ.
- $\eta(\frac{\pi}{2}) = 0$
- $\eta(0) = 1$

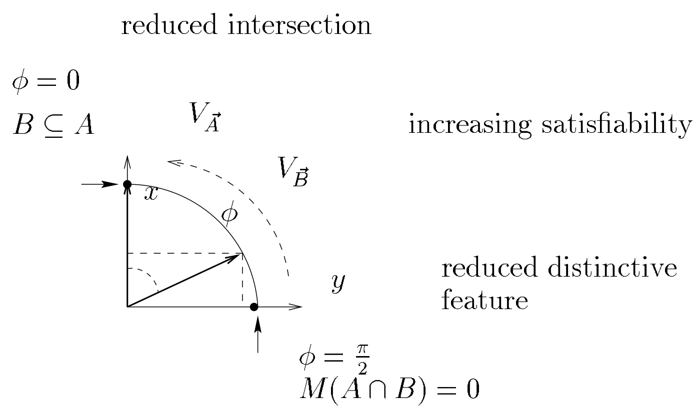

Figure 2: New representation of a measure of satisfiability

This new form of a measure of satisfiablity, expressed by a unique variable, has the advantage not to be dependent upon the size of the system. Furthermore, this normalization makes the definition of a measure of satisfiability more simple insofar as the argument is a segment $[0, \frac{\pi}{2}]$ and not a quarter of plan.

There are of course many possible choices for the satisfiability measure η satisfying these three conditions. Among them, let us distinguish the two following forms:

- $\eta_1(\phi) = 1 - \frac{2}{\pi}\phi$. It is the linear satisfiability function.
- $\eta_2(\phi) = \cos\phi$. This function has the advantage of presenting a meaningful physical insight. If we represent the reference set A by the

$(1,0)$ vector $V_{\vec{A}}$ in figure 2. If we describe each set B and its related point (x,y) by a vector $V_{\vec{B}}$ from the origin, $\eta_2(\phi)$ is the scalar product $V_{\vec{A}} \cdot V_{\vec{B}}$. When the two vectors are orthogonal, then the satisfiability vanishes: $S(A,B) = 0$. This is a good signification of orthogonality. More generally, the satisfiability appears as a projection, and the lack of satisfiability is represented as a deviation in figure 2: this is an intuitive notion of satisfiability.

We can also focus on the following measures which are known in the literature:

- $S(A,B) = \frac{M(A \cap B)}{M(B)}$, which is usually defined with M the sigma-count, can be also written: $\eta_3(\phi) = \frac{1}{1+\tan\phi}$.
- $S(A,B) = 1 - M(B-A)$ was introduced in [4] with $M(A) = \sup_x f_A(x)$ and with the difference $f_{A-B}(x) = \begin{cases} f_A(x) & \text{if } f_B(x) = 0 \\ 0 & \text{if } f_B(x) > 0 \end{cases}$. It can be also written: $\eta_4(\phi) = 1 - \sin\phi$.

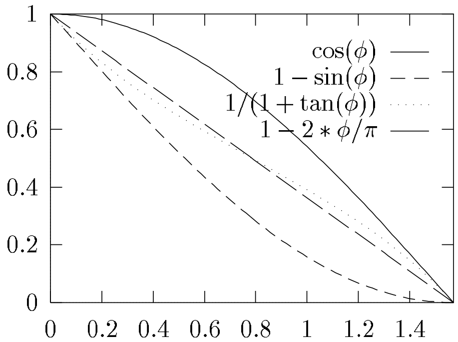

Figure 3: Satisfiability measures

Figure 3 displays the behaviour of various mentioned measures of satisfiability. We can see that:

- η_1 is linearly discriminant: satisfiability decreases linearly with the deviation.
- η_2 is discriminant for low satisfiability: a small difference between a set and the reference is tolerated.

- η_4 is discriminant for high satisfiablity: a small difference between a set and the reference is not tolerated.

  Remark η_4 and η_2 are symmetrical relatively to $(\frac{\pi}{4}, \frac{1}{2})$: $\eta_4(\frac{\pi}{2}-\phi) = 1-\eta_2(\phi)$. η_2 has a low discrimination power for high satisfiability, and a high discrimination power for poor satisfiability, whereas η_4 does exactly the reverse.

- η_3 is discriminant for high and low satisfiability at the same time. For high satisfiability, this measure is between the linearly discriminant measure and the low satisfiability discriminant measure. For low satisfiability, it is between the linearly discriminant measure and the high satisfiability discriminant measure. But it is low discriminant for $\eta(\phi) = 1/2$.

We can consider that the *discrimination power* of a measure of satisfiability is given by the derivative $\eta'(\phi)$ of η. For instance, for small and large differences between a set and the reference that is to say for $\phi = 0$ or $\phi = \frac{\pi}{2}$:

- $\eta_1'(\phi) = -\frac{2}{\pi}$, for all ϕ.
- $\eta_2'(0) = 0$ and $\eta_2'(\frac{\pi}{2}) = -1$
- $\eta_3'(0) = -1$ and $\eta_3'(\frac{\pi}{2}) = -1$
- $\eta_4'(0) = -1$ and $\eta_4'(\frac{\pi}{2}) = 0$

In general, for every possible η, we have:

$$\int_0^{\frac{\pi}{2}} \eta'(\phi)d\phi = \eta(\frac{\pi}{2}) - \eta(0) = -1$$

This means that the total discrimination power $\eta'(\phi)$ has to be distributed on the $[0, \frac{\pi}{2}]$ interval, but a high discrimination power somewhere implies a low discrimination power elsewhere, the integral being constant.

Accordingly, it is necessary to choose a measure with a discrimination power suitable for the considered application. This suggests a method of construction of a measure of satisfiability. The choice of the discrimination power is the first step. Then η is obtained by integration of this function $\eta'(\phi)$.

For instance, a function with a high discrimination power for $\eta(\phi) = 1/2$ but a low discrimination for $\eta(\phi) = 0$ and $\eta(\phi) = \frac{\pi}{2}$ is needed. This kind of measures means that if a description is not far from the reference, then the satisfiability is near from 1 because the difference is not significative.

If a description is very far from the reference, we can consider that the satisfiability is null. The example of this method of self-construction is illustrated in figure 4.

Once the behaviour of the wanted measure is known, the measure η can be computed thanks to the discrimination power η'. For instance, the discrimination power can be considered:

- constant by intervals. In this case, η is piecewise linear (see figure 4).

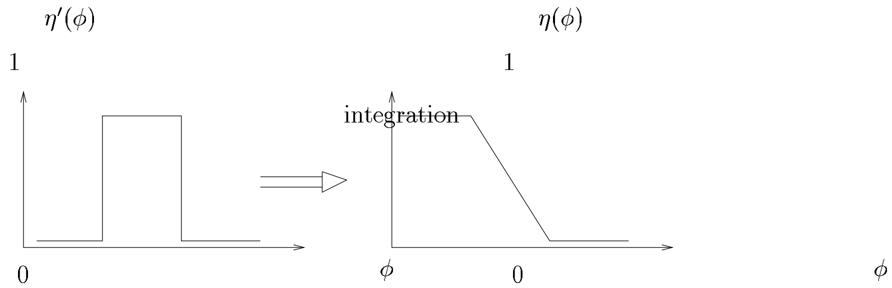

Figure 4: Construction of a measure of satisfiability with a constant by intervals discrimination power

  This rigid shape of function is not very satisfactory for a measure of satisfiability.

- A better measure is derived from a discrimination power with a more gradual form. Let us give two possible solutions:

  1. the Lorentz's function: $\eta'(\phi) = \frac{1}{1+\frac{(\phi-\phi_0)^2}{\Gamma^2}}$
  2. the Gaussian function: $\eta'(\phi) = \exp^{-\frac{(\phi-\phi_0)^2}{\Gamma^2}}$

  normalized on $[0, \frac{\pi}{2}]$ (see figure 5). Γ is a parameter which allows a fine control.

Finally, we would like to give the exemple of an interesting function for the measure of satisfiability using the Fermi-Dirac function. The analytic form is:

$$F_{FD}(\phi) = \frac{1}{1+\exp^{\frac{(\phi-\frac{\pi}{4})}{\Gamma}}}$$

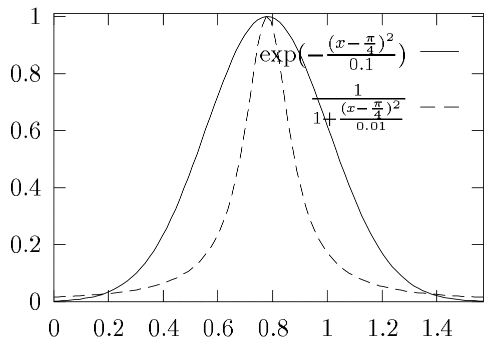

Figure 5: Discrimination power defined by Lorentz and Gaussian functions

and

$$\eta(\phi) = \frac{F_{FD}(\phi) - F_{FD}(\frac{\pi}{2})}{F_{FD}(0) - F_{FD}(\frac{\pi}{2})}$$

$\eta(\phi)$ reflects essentially the Fermi-Dirac function. The above expression is used to ensure that $\eta(0) = 1$ and $\eta(\frac{\pi}{2}) = 0$.

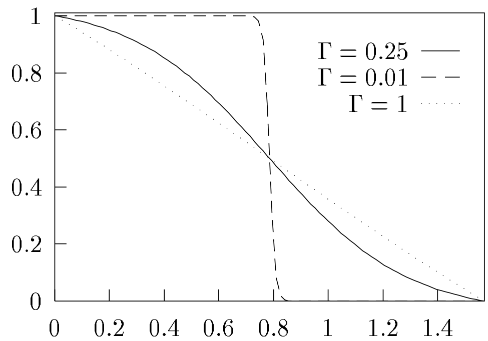

Figure 6: The Fermi-Dirac measure of satisfiability

The interest of this function lies on its physical meaning: in a physical system of temperature T, containing a statistical set of states, the Fermi-Dirac function describes the statistical probability that a state of energy ϕ is filled or not, with $\Gamma = kT$, where k is the Boltzmann constant. Hence, it describes the probability of a state to belong to the Fermi sea [5] which is a good illustration of satisfiability. Γ controls the decrease of the curve. The choice of Γ enables to define a measure of satisfiability more or less severe, as shown on figure 6.

2.2 Measures of resemblance

We are now interested in resemblance measures.

A measure of *resemblance* is used for a comparison between the descriptions of two objects, of the same level of generality, to decide if they have many common characteristics.

Measures of resemblance are appropriate for a case-based reasoning or an instance-based learning. In clustering methods, distances can be replaced by a measure of resemblance. More generally, similarity-based classification methods [10], [11] have to use resemblance measures as soon as all objects have the same level of generality.

Definition 5 *An M-*measure of resemblance *on Ω is a measure $S(A,B) = F_S(X,Y,Z)$ such that:*

- *F_S is increasing with X and decreasing with Y and Z*
- *$F_S(X,0,0) = 1$ for all X*
- *$F_S(X,Y,Z) = F_S(X,Z,Y)$.*

M-measures of resemblance which satisfy an additional property of t-transitivity, for a triangular norm t, are extensions of indistinguishability relations [13], [15] to fuzzy sets. In the case where t is the minimum, we obtain extensions of measures of similarity.

M-measures of resemblance satisfying the property of exclusiveness:

$$F_S(0,Y,Z) = 0 \quad \text{for all} \quad (Y,Z) \neq (0,0)$$

are called *exclusive M-measures of resemblance.* We focus on them in the sequel.

Following our normalization procedure, we define:

$$x = \frac{X}{\sqrt{X^2+Y^2+Z^2}}$$
$$y = \frac{Y}{\sqrt{X^2+Y^2+Z^2}}$$
$$z = \frac{Z}{\sqrt{X^2+Y^2+Z^2}}$$

for $(X, Y, Z) \neq (0,0,0)$. Similarly to the the case of measures of satisfiability, this ensures that an exclusive measure of resemblance is not dependent on the scale of the problem.

The domain of study is now restricted to a piece of the unity sphere since $x^2 + y^2 + z^2 = 1$. We have now to find the exclusive resemblance μ which satisfies the symmetry property $\mu(x, y, z) = \mu(x, z, y)$.

Geometrically, the sphere is simply obtained by a rotation of the satisfiability circle around the x-axis (see figure 7). The vector representation is still valid.

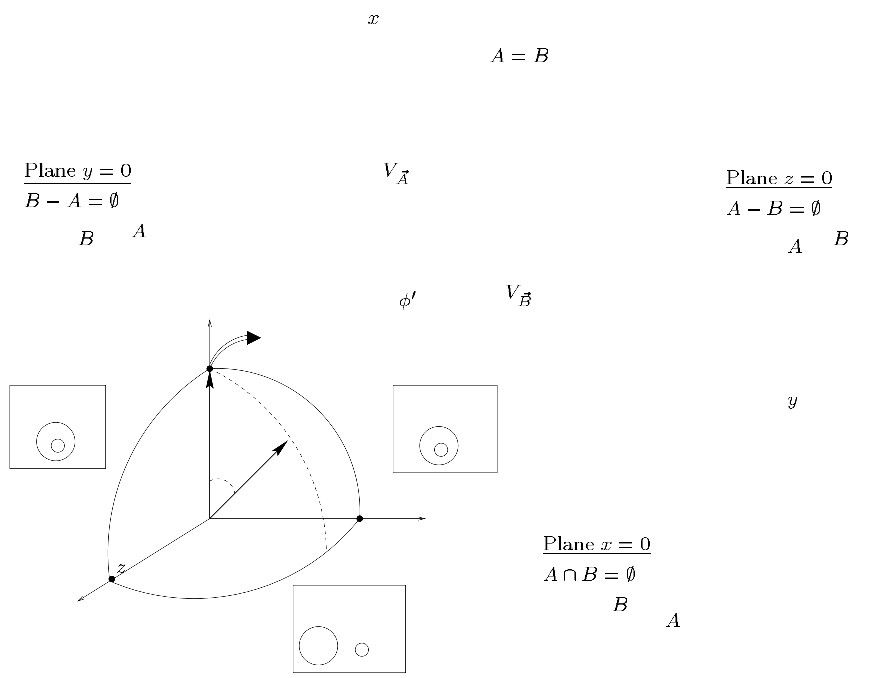

Figure 7: New representation of an exclusive measure of resemblance

Let us consider $\rho = \xi(y, z)$ with $\xi(y, z) = \xi(z, y)$. This means that ρ can be described by any symmetrical function with respect to y and z.

Let us look for an expression of $\nu(x, \rho) = \mu(x, y, z)$. By definition, an exclusive measure of resemblance is symmetrical with respect to (Y, Z), (see

Definition 5). And we have:

$$\begin{aligned} \nu(0,\rho) &= 0 \text{ if } \rho \neq 0 \\ \nu(x,0) &= 1 \ \forall\, x \end{aligned}$$

These conditions show that the problem has been reduced to a satisfiability measure. We can therefore use again the solution described in the preceding section dealing with satisfiability. With this definition of ρ, *an exclusive resemblance appears as a satisfiability where a global distinctive feature ρ is defined by $\rho = \xi(y,z)$*, from the two individual distinctive features y and z (see Figure 8).

Satisfiability		Resemblance
(x,y)	$\longrightarrow$	(x,ρ) with $\rho = \xi(y,z)$

Figure 8: Correspondance between saistisfiability measures and exclusive resemblance measures

We can also consider different exclusive measures of resemblance as we have already done with measures of satisfiability. Let us define $\psi = \arctan(\frac{\rho^0}{x})$ with $\rho^0 = y + z$.

- $\nu_1 = \frac{1}{1+\frac{\rho^0}{x}} = \frac{1}{1+\tan\psi}$. This measure corresponds to the measure of satisfiability η_3. Furthermore, ν_1 can be also written as: $S(A,B) = \frac{M(A\cap B)}{M(A\cup B)}$ with M such that: $M(A\cup B) = M(A\cap B) + M(A-B) + M(B-A)$. This measure was introduced in [8].

Other definitions of ρ can be envisaged, for instance:

$$\begin{aligned} \rho' &= \sqrt{y^2+z^2} \\ \text{or } \rho'' &= (\sqrt{y}+\sqrt{z})^2 \end{aligned}$$

The choice of a particular form of ρ has an effect on the measure of resemblance because this parameter represents distinctive elements. We can notice that:

$$\rho'' \geq \rho^0 \geq \rho' \ \forall\, y,z \tag{1}$$

As ρ has a decreasing effect on an exclusive measure of resemblance, relation (1) implies that:

$$\text{for a given } x \text{ and for all } y \text{ and } z,\ \nu''(x,\rho'') \leq \nu(x,\rho^0) \leq \nu'(x,\rho') \qquad (2)$$

Relation (2) means that $\nu''(x,\rho'')$ penalizes more the differences between two sets than $\nu(x,\rho^0)$ and that $\nu(x,\rho^0)$ penalizes more the differences than $\nu'(x,\rho')$. Furthermore, a particular ρ is sensitive to the symmetry between y and z as it is illustrated in Figure 9. Indeed, if differences are unbalanced, it means that $y >> z$ or $z >> y$, or inversely, if differences are balanced, it means that $y \approx z$, the behaviours of a given ρ are not the same.

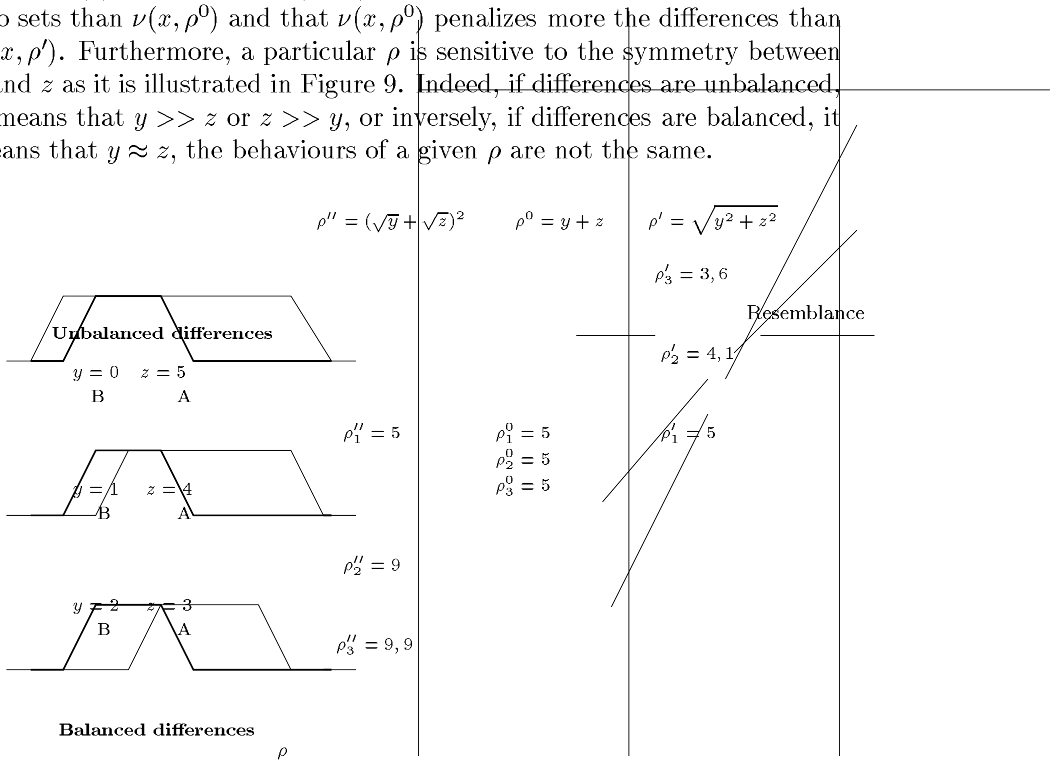

Figure 9: The effects of different definitions of ρ on exclusive measures of resemblance.

3 Conclusion

This paper gives an explicit method to choose a measure of comparison. This methods consists in two major points:

1. the choice of the appropriate family of measures of comparison.

2. the choice of the appropriate measure in the chosen family of measures of comparison.

The first choice was described in details in [3]. This paper focuses more on the second point. It establishes the notion of *discrimination power* which enables to describe precisely the behaviour of a measure. Therefore, it is possible to choose a particular measure among measures of same family that the chosen one because of its particular power of discrimination.

References

[1] S. Bothorel, B. Bouchon, and S. Muller. A fuzzy logic-based approach for semiological analysis of microcalcification in mammographic images. *International Journal of Intelligent Systems*, 1997, to appear.

[2] B. Bouchon-Meunier, C. Marsala, and M. Ramdani. Learning from imperfect data. In H. Prade D. Dubois and R. R. Yager, editors, *Fuzzy Information Engineering: a Guided Tour of Applications*, pages 139–148. John Wileys and Sons, 1997.

[3] B. Bouchon-Meunier, M. Rifqi, and S. Bothorel. Towards general measures of comparison of objects. *Fuzzy Sets and Systems*, 84(2):143–153, 1996.

[4] B. Bouchon-Meunier and L. Valverde. Analogy relations and inference. In *Proceedings of* 2^{nd} *IEEE International Conference on Fuzzy Systems*, pages 1140–1144, San Fransisco, 1993.

[5] Berkeley Physics Course. Statistical physics. vol. 5.

[6] E. Diday. Introduction à la méthode des nuées dynamiques. In *Analyse des données*, volume I, pages 121–132. A. P. M. E. P (Association des Professeurs de Mathématiques de l'Enseignement Public), 1980.

[7] E. Diday, J. Lemaire, J. Pouget, and F. Testu. *Eléments d'analyse de données*. Dunod, 1982.

[8] D. Dubois and H. Prade. *Fuzzy Sets and Systems, Theory and Applications*. Academic Press, New- York, 1980.

[9] D. Dubois and H. Prade. A unifying view of comparison indices in a fuzzy set-theoretic framework. In R. R. Yager, editor, *Fuzzy set and possibility theory*, pages 3–13. Pergamon Press, 1982.

[10] M. Rifqi. Constructing prototypes from large databases. In *IPMU'96*, pages 301–306, Granada, 1996.

[11] M. Rifqi, S. Bothorel, B. Bouchon-Meunier, and S. Muller. Similarity and prototype based approach for classification of microcalcifications. In 7^{th} *IFSA World Congress*, pages 123–128, Prague, 1997.

[12] K. Shiina. A fuzzy-set-theoretic feature model and its application to asymmetric data analysis. *Japanese psychological research*, 30(3):95–104, 1988.

[13] E. Trillas and L. Valverde. On implication and indistinguishability in the setting of fuzzy logic. In J. Kacprzyk and R. R. Yager, editors, *Management Decision Support Systems Using Fuzzy Sets and Possibility Theory*. Verlag TUV, Rheinland, 1984.

[14] A. Tversky. Features of similarity. *Psychological Review*, 84:327–352, 1977.

[15] L. Valverde. On the structure of t-indistinguishability operators. *Fuzzy Sets and Systems*, 17:313–328, 1985.