A CLASSIFICATION METHOD BASED ON INDISTINGUISHABILITIES

Luis Garmendia
Technical University of Madrid.
E.T.S.I. Caminos
Dpto. de Matemática Aplicada.
and IBM Learning Services.
luisgarmendia@es.ibm.com

Adela Salvador
Technical University of Madrid.
E.T.S.I. Caminos
Dpto. de Matemática Aplicada
E-mail: ma09@caminos.upm.es

Enric Trillas
Technical University of Madrid.
Facultad de Informática.
Dpto. de Inteligencia Artificial
E-mail: etrillas@fi.upm.es

Abstract:

This paper works on the inference independent clustering method to solve the problem of classifying a classical set given a fuzzy set and a T-indistinguishability.

It is shown an application of this method for computing measures of specificity of fuzzy sets under T-indistinguishabilities.

Keywords: T-indistinguishability, inference independent classes, clustering.

1 INTRODUCTION

This paper gives a method to classify in classes a finite set X given a fuzzy set and a T-indistinguishability on X.

The new method builds up a set of classes of X that are 'inference independent', that is, a set of classes in such a way that given two elements of X in different classes, it is not possible to infer a degree of membership of an element greater than the degree of membership given by μ by fuzzy inference through the t-norm T and the T-indistinguishability with the other element.

It is also given a fuzzy set on the set of inference independent classes.

It is shown an application of the method to compute measures of specificity of fuzzy sets under T-indistinguishabilities. When the knowledge available is increased through a T-indistinguishability, the specificity of fuzzy sets is also increased. The specificity of a fuzzy set under a T-indistinguishability can be computed as the specificity of the before defined fuzzy set.

X will be a crisp finite set, and $R: X \times X \to [0, 1]$ a **T-indistinguishability** (that is, R is reflexive, simetric and T-transitive).

A T-indistinguishability is called a similarity when T = mín.

2 INFERENCE INDEPENDENT CLASSES

Let μ be a fuzzy set on a finite space $X = \{x_1, ..., x_n\}$, let S be a T-indistinguishability on X and let T be a t-norm.

Definition

x_k is related with x_j, and it is denoted by $\mathbf{x_k \succeq x_j}$, if and only if $T(\mu(x_k), S(x_k, x_j)) \geq \mu(x_j)$.

Two elements x_k and x_j in X are in the same inference independent class if and only if they are comparable by the $\succeq$ preorder.

So, it is defined the class of an element x_k as follows:

$\mathbf{[x_k] = \{x_j \text{ such that } x_k \succeq x_j \text{ or } x_j \succeq x_k\}}$

This definition means that when x_k is related with x_j, it is possible to deduce the same or more of what we know of x_j from x_k by making fuzzy inference with the t-norm T ant the given T-indistinguishability. That is, x_k is related with x_j when the information on x_j is increased by knowing the membership degree of x_k and its T-indistinguishability relation with x_j.

Proposition

Let μ be a fuzzy set on a finite set $X=\{x_1, ..., x_n\}$ and let S be a T-indistinguishability, then the relation $\succeq$ is a classical preorder relation on X.

Proof

- The $\succeq$ relation is reflexive:

$T(\mu(x_i), S(x_i, x_i))) = T(\mu(x_i), 1) = \mu(x_i)$, so $x_i \succeq x_i$.

- The $\succeq$ relation is transitive:

Let´s suppose that $x_i \succeq x_j$ and $x_j \succeq x_k$.
$x_i \succeq x_j$, so $T(\mu(x_i), S(x_i, x_j)) \geq \mu(x_j)$.
$x_j \succeq x_k$, so $T(\mu(x_j), S(x_j, x_k)) \geq \mu(x_k)$.
Hence $\mu(x_k) \leq T(\mu(x_j), S(x_j, x_k))$
$\leq T(T(\mu(x_i), S(x_i, x_j)), S(x_j, x_k))$ (T is associative)
$= T(\mu(x_i), T(S(x_i, x_j), S(x_j, x_k))$ (S is T-transitive)
$\leq T(\mu(x_i), S(x_j, x_k))$
and so $x_i \succeq x_k$. ∎

Definition
The fuzzy set $\Im$ on the crisp set of inference independent classes is defined as follows:
$\Im([x_k]) = \mathbf{Max_j(T(\mu(x_j), S(x_k, x_j)))}$.

It is trivial to show that the membership degree of $[x_k]$ in $\Im$ is greater or equal than $\mu(x_k)$ for all x_k in X.

3 ALGORITHM TO COMPUTE INFERENCE INDEPENDENT CLASSES

By the previous definition, the membership degree of the classes of the elements $\{x_1, ..., x_n\}$ is computed by the Max-T rule of compositional inference using the fuzzy set μ and the T-indistinguishability S.
So, $\Im([x_j]) = Max_k(T(\mu(x_k), S(x_k, x_j)))$
$= T(\mu(x_k), S(x_k, x_j))$ for a particular k.
If $k\neq j$, then x_k represents the dass of x_j ($x_k \succeq x_j$).
The following algorithm´s purpose is to get rid of elements x_j when it exits a k such that $[x_j] = [x_k]$ because $x_k \succeq x_j$.
The transitive property of the $\succeq$ relation is necesary for this algorithm to finish in a few steps, because when an element x_k represents another element x_j ($x_k \succeq x_j$), we can eliminate x_j without taking care that x_j could represent a third element x_i ($x_j \succeq x_i$), because in this case x_k would represent x_i ($x_k \succeq x_i$) and x_i, x_j and x_k would belong to $[x_k]$. As x_k would represent x_i, the algorithm also gets rid of x_i in a further step.
In summary, the algorithm detects and eliminates the elements of X that are represented by other elements, toward getting a final set of elements X' that represents the different inference independent classes ($X'\subseteq X$).

This algorithm steps are the following:

Step 1: Compute μ o$_{Max-T}$ S(*, x_1).

If $Max_j(T(\mu(x_j), S(x_j, x_1))) \geq \mu(x_1)$ for some $j\neq 1$ then $X^1:=X-\{x_1\}$, and μ^1 and S^1 are the restrictions of de μ and S to X^1.

Otherwise, $X^1 = X$ (x_1 represents its own class and will belong to the final set of classes X').

Step k: Compute μ o$_{Max-T}$ S(*, x_k).
If $Max_j(T(\mu^{k-1}(x_j), S^{k-1}(x_j, x_k))) \geq \mu^{k-1}(x_k)$ for some $j \neq k$ then $X^k=X^{k-1} - \{x_k\}$.
Otherwise $X^k = X^{k-1}$.

Repeating this process until the n[th] step, the set X^n = X' is the set of inference independent classes and their membership degree are given by the fuzzy set restricted to X^n

Example
Let μ be the fuzzy set on $X=\{x_1, .., x_5\}$:
$\mu = 1/x_1+0.7/x_2+ 0.5/x_3+0.2/x_4+ 0/x_5$.
Let S be a T-Indistinguishability represented by

$$S = \begin{pmatrix} 1 & 1 & 0 & 0.5 & 0.2 \\ 1 & 1 & 0 & 0.5 & 0.2 \\ 0 & 0 & 1 & 0 & 0 \\ 0.5 & 0.5 & 0 & 1 & 0.2 \\ 0.2 & 0.2 & 0 & 0.2 & 1 \end{pmatrix},$$

which is reflexive and Min-transitive.
Let T be the t-norm minimum. The membership degree $\Im$ of the inference independent classes for the elements x_i are the following:

$\mu \, o_{Max\text{-}Min} \, S$

$= (1, 0.7, 0.5, 0.2, 0) \, o_{Max\text{-}Min} \, S$

$= (1, 1, 0.5, 0.5, 0.2)$

The following algorithm steps are done to decide a set of classes that are Min-inference independent.

Step 1 Compute

$\mu \, o_{Max\text{-}Min} \, S(*, x_1) =$

$$(1, 0.7, 0.5, 0.2, 0) \, o_{Max\text{-}Min} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0.5 \\ 0.2 \end{pmatrix}$$

$= Max\{1, 0.7, 0, 0.2, 0\} = 1 = \mu(x_1)$.

As $1 = Max(T(\mu(x_1), S(x_1, x_1))) = \mu(x_1)$, then x_1 represents its own class, and $[x_1]$ belongs to $X / \succeq$.
So $X^1=X$, $\mu^1= \mu$ and $S^1=S$.

Step 2 Compute

$\mu^1 \, o_{Max\text{-}T} \, S^1(*, x_2) =$

$$(1, 0.7, 0.5, 0.2, 0) \, o_{Max\text{-}Min} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0.5 \\ 0.2 \end{pmatrix} =$$

$Max\{1, 0.7, 0, 0.2, 0\} = 1 \geq 0.7 = \mu(x_2)$.

As $1 = Max(T(\mu^1(x_1), S^1(x_1, x_2))) \geq \mu^1(x_2)$ then x_1 represents x_2 (by the relation $\succeq$), that is, $[x_1] = [x_2]$, so

$X^2 = X^1\text{-}\{x_2\} = \{x_1, x_3, x_4, x_5\}$, μ^2 is μ^1 restricted to X^1 and

$$S^2 = \begin{pmatrix} 1 & 0 & 0.5 & 0.2 \\ 0 & 1 & 0 & 0 \\ 0.5 & 0 & 1 & 0.2 \\ 0.2 & 0 & 0.2 & 1 \end{pmatrix}$$

on $X^2 \times X^2$, (that is, on $\{x_1, x_3, x_4, x_5\}^2$).

Step 3 Compute

$\mu^2 \, o_{Max\text{-}T} \, S^2(*, x_3) =$

$$(1, 0.5, 0.2, 0) \, o_{Max\text{-}Min} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} =$$

$Max\{0, 0.5, 0, 0\} = 0.5 = \mu(x_3)$.

As $0.5 = \mu(x_3)$, x_3 represents its own class and $[x_3]$ will be a new element in $X / \succeq$.
$X^3 = X^2$, $\mu^3 = \mu^2$ and $S^3=S^2$.

Step 4 Compute

$\mu^3 \, o_{Max\text{-}T} \, S^3(*, x_4) =$

$$(1, 0.5, 0.2, 0) \, o_{Max\text{-}Min} \begin{pmatrix} 0.5 \\ 0 \\ 1 \\ 0.2 \end{pmatrix} =$$

$Max\{0.5, 0, 0.2, 0\} = 0.5 \geq 0.2 = \mu(x_4)$.

As $0.5 = Max(T(\mu^3(x_1), S^3(x_1, x_4))) \geq \mu^3(x_4) = 0.2$ then x_1 represents x_4 by the relation $\succeq$, so
$X^4 = X^3\text{-}\{x_4\} = \{x_1, x_3, x_5\}$, μ^4 is μ^3 restricted to X^4 and

$$S^4 = \begin{pmatrix} 1 & 0 & 0.2 \\ 0 & 1 & 0 \\ 0.2 & 0 & 1 \end{pmatrix}$$

on $X^4 \times X^4$.

Step 5 computes

$\mu^4 \, o_{Max\text{-}T} \, S^4(*, x_5) =$

$$(1, 0.5, 0) \, o_{Max\text{-}Min} \begin{pmatrix} 0.2 \\ 0 \\ 1 \end{pmatrix} =$$

$Max\{0.2, 0, 0\} = 0.2 \geq 0 = \mu(x_5)$.

As $0.2 = Max(T(\mu^3(x_1), S^4(x_1, x_5))) \geq \mu^4(x_5) = 0$ then x_1 represents x_5 by the relation $\succeq$.
$X^5 = X^4 - \{x_5\} = \{x_1, x_3\}$, so the set of Min-inference independent classes in $X / \succeq$ are $\{[x_1], [x_3]\} = \{\{x_1, x_2, x_4, x_5\}, \{x_3\}\}$, and their membership degree are those of $\Im$ restricted to X^5, that is, $\{1/x_1, 0.5/x_3\}$.

4 APPLYING THE ALGORITHM TO COMPUTE SOME MEASURES OF SPECIFICITY.

Yager [Yager; 1991, 91] introduced the concept of specificity of a fuzzy set under similarities using the *Zadeh* [Zadeh; 1971] concept of similarity or Min-indistinguishability.
The α-cut of a similarity S is a classical equivalence relation [3] denoted S_α. Let π_α be the set of equivalence classes of S for a given value α. Let μ_α/S be the set of equivalence classes of π_α defined in the following way: class $\pi_\alpha(i)$ belongs to μ_α/S if there exists an element x contained in $\pi_\alpha(i)$ and in the μ´s α-cut (μ_α).

Definition
Yager [1991, 91] definition of measure of specificity of a fuzzy set μ under a similarity is the following:

$$\mathbf{S_p(m/S)} = \int_0^{\alpha_{max}} \frac{1}{Card(\mu_\alpha / S)} \, \mathbf{d}a \, .$$

The measure of specificity under similarities are maximal when μ_α is contained in one class of S_α for all α.

This definition is good enough when the information is increased by a similarity, but it is not well defined for any T-indistinguishability, because when T is not the minimun t-norm the α-cut of S is not an equivalence relation and then μ_α/S is not well defined.

Definition
Let Sp be a measure of specificity.
A **measure of specificity of a fuzzy set** m **under a T-indistinguishability** S is the measure of specificity Sp of the fuzzy set $\Im$ on the set of classes $X^n = X / \succeq$.

Theorem
The measure of specificity of μ under S computed by the algorithm satisfies the four axioms of a measure of specificity under a T-indistinguishability.

Proof
The proof is in [1].

EXAMPLE
When using the previous example given to show how the algorithm works, the following Min-inference independent classes are found:
$\{\{x_1, x_2, x_4, x_5\}, \{x_3\}\}$. Their membership degrees to the fuzzy set $\Im$ are $1/[x_1]+0.5/[x_3]$.
So, **the measure of specificity of** m **under the Min-indistinguishability S is the measure of specificity of the fuzzy set** Á on the set of classes $[x_1]$ and $[x_3]$ with membership degrees $1/[x_1]+ 0.5/[x_3]$.
When using, for example, the linear measure of specificity of Yager [1990] with a weight $w_2 = 1$, the measure of specificity of μ under S is:

$$\begin{aligned} &Sp(\mu / S) \\ &= Sp(\Im) \\ &= Sp(1/[x_1]+0.5/[x_3]) \\ &= 1 - 0.5 = 0.5. \end{aligned}$$

Compare this result with the linear measure of specificity of μ with a weight $w_2 = 1$.

$$\begin{aligned} &Sp(\mu) \\ &= Sp(1/x_1+0.7/x_2+ 0.5/x_3+0.2/x_4+ 0/x_5) \\ &= 1 - 0.7 = 0.3. \end{aligned}$$

Observe that the measure of specificity of μ under a T-indistinguishability S is greater than the measure of specificity of μ. This is because the T-indistinguishability adds information which tells that four of the five elemetns of X are similar. When using the measure of specificity as a measure of the information of a fuzzy set in order to make a decision of an element of X, the Min-indistinguishability is telling us that four of the five possible decisions are similar, so the decision is simplify to two classes of elements of X.

The fuzzy set $\Im$ on the inference independent classes is usefull to define and compute new measures of specificity of a fuzzy set μ when the information is increased by a T-indistinguishability.

Conclusions

This paper gives an algorithm to classify a finite set X into inference independent classes given a fuzzy set and a T-indistinguishability on X.

A fuzzy set $\Im$ on the set of inference independent classes is given.

It is shown an application of the algorithm to define and compute new measures of specificity of a fuzzy set under a T-indistinguishability.

Adknoledgements

This paper has been partially subventioned by the spanish proyects CICYT: PB-98-1379-c02-02 (MEC) and TIC 00-1420.

References

[1] L. Garmendia. Contribución al Estudio de las Medidas en la Lógica Borrosa: Condicionalidad, Especificidad and Transitividad. Madrid. 2001. Ph. D. Thesis. Universidad Politécnica de Madrid. E.T.S.I. Caminos, Canales y Puertos (in Spanish).

[2] L. Garmendia, R.R. Yager, E. Trillas, A. Salvador. On t-norms based measures of specificity. *Fuzzy Sets and Systems.* Acepted for publication in 2000.

[3] E. Trillas, C. Alsina and J. M. Terricabras. *Introducción a la Lógica Borrosa*. Editorial Ariel. Barcelona (in Spanish). 1995.

[4] R. R Yager, Ordinal measures of specificity. *International Journal of General Systems* 17, pp. 57-72, 1990.

[5] R. R Yager, Similarity based measures of specificity. *International Journal of General Systems* 19, pp. 91-106, 1991.

[6] L. A. Zadeh, Similarity relations and fuzzy orderings. *Information Sciences* 3, pp. 177-200, 1971.