A Characterization of Linearly Compensated Hybrid Connectives Used in Fuzzy Classifications

Mónica Sánchez[1] and **Francesc Prats**[1] and **Núria Agell**[2] and **Joseph Aguilar-Martin**[3]

Abstract. The study of linearly compensated hybrid connectives $H = \lambda C + (1-\lambda)\bar{C}$, where C is a t-norm and $\bar{C}$ represents the dual connective of C, to define aggregation operators for fuzzy classifications is a key point not only in fuzzy sets theory but also in learning processes. Although these operators are not associative, the fact that they can be decomposed into associative functions easily gives rise to n-ary aggregation functions by straightforward iteration. Among the most commonly used t-norms are those of Frank's family, which are simultaneously t-norms and copulas. The purpose of this paper is to give a characterization of the hybrid connective H, via the properties of the connective C. Necessary and sufficient conditions of H that define C as a copula are given. The characterized hybrid connectives H are used to compute the global adequacy degree of an object in a class from marginal adequacy degrees in a learning system.

Keywords. Hybrid Connectives, Reasoning under Uncertainty, Machine Learning, Classification Algorithms, Qualitative Reasoning.

1 INTRODUCTION

There are many ways to recognize an object as a member of a given class. One of them is based on the concept of similarity between objects, within a mathematical structure of binary relations in a set. Another perhaps more primitive concept is the adequacy of an object in a class, which can be treated as the membership function of a particular fuzzy set [4],[1].

Within this context the study of linearly compensated hybrid connectives to perform the aggregation of possibility functions will allow the epistemological attitude of the observer to be captured. In this respect, fuzzy logic represents a mathematical tool for the modulation of the different human attitudes, usually half way between two extremes, from the strongest exigency to the loosest requirement. The former corresponds to the simultaneous conjunction, and the latter is associated with total disjunction.

This concept of linear compensation between diverse human attitudes was implemented in LAMDA (Learning Algorithm for Multivariate Data Analysis) [4]. This learning algorithm is an incremental conceptual clustering system based on the heuristic rule of maximum adequacy to form classes, which has obtained good results in several fields, such as segmentation of customer profiles in marketing. It computes a global adequacy degree between the object and any of the classes, and it assigns such object to the class with the maximum adequacy degree. One of the advantages of this learning algorithm, which is crucial in marketing studies, is that the adequacy of any object to every class is computed and that it can be used to analyze the overall performance of the population [1].

[1] Technical University of Catalonia, Barcelona, Spain email: francesc.prats@upc.es,monica.sanchez@upc.es
[2] Esade, Ramon Llull University, Barcelona, Spain email:agell@esade.edu
[3] LAAS-CNRS, Toulouse, France email:aguilar@laas.fr

The global adequacy degree of an object in a class is computed by the "connection" between the partial information or its marginal adequacy degrees [2]. This generates a fuzzy partition on the initial set, which will depend on the choice of the connectives used to operate the fuzzy sets corresponding to marginal adequacy degrees. LAMDA methodology uses a kind of hybrid connectives in its learning process, which are linearly compensated between the conjunction and the disjunction.

The aim of this paper is to present a characterization of the hybrid connective $H = \lambda C + (1-\lambda)\bar{C}$ obtained via linear convex combinations between C and its dual operator, via the properties of the connective C.

2 LINEARLY COMPENSATED HYBRID CONNECTIVES

From the most exigent point of view, two objects with at least one different descriptor value are likely to be regarded as different and assigned to different classes. This approach is associated with the conjunction, or intersection, represented in fuzzy logic by t-norms. The least exigent approach consists of considering that the objects having at least one descriptor with the same value are likely to be granted the same assignation. This is associated with the disjunction, or union, represented in fuzzy logic by t-conorms, which are the dual functions of t-norms [3].

LAMDA uses the convex linear combination, by means of a parameter $\lambda \in [0,1]$, of a special kind of t-norms (those of Frank's family [3]) and their dual t-conorms. It should be pointed out that connectives in Frank's family are simultaneously t-norms and n-copulas and that n-copulas are important connectives in Probability Theory since every n-dimensional distribution function is of the form $C(F_1, \ldots, F_n)$, where C is a n-copula and $F_1, \ldots, F_n$ are margin functions. This family includes the most commonly used t-norms, such as Min, Product or the Lukasievic t-norm. The hybrid connectives considered in this work are linear convex combinations between a 2-copula and its dual:

Definition 1 *The* hybrid connective H associated with a 2-copula C and a real number $\lambda \in [0,1]$ *is the binary operation on* $[0,1]$ *defined by*

$$H = \lambda C + (1-\lambda)\overline{C}, \tag{1}$$

where $\overline{C}$ is the dual operator of C. The parameter λ is named *tolerance parameter* and allows a modulation of exigency-tolerance from $\lambda = 0$ which corresponds to the case $H = C$ until $\lambda = 1$ which corresponds to the case $H = \overline{C}$. For a given situation with a pre-established number of classes and for a finite and fixed database, the number of unrecognized elements is directly linked to the exigency and therefore to the choice of λ. By continuously varying $\lambda \in [0,1]$, all possible partitions can be obtained within a family of chosen connectives and therefore a criterion for the choice of λ.

3 A CHARACTERIZATION OF HYBRID CONNECTIVES

This section provides a characterization of the hybrid connectives

$$H = \lambda C + (1 - \lambda)\overline{C} \quad (2)$$

where $\overline{C}$ is the dual connective of C (as in definition 6) and $\lambda \in [0, 1]$.

Lemma 1 *Let $\lambda \neq 1/2$ and $\overline{C}(x, y) = x + y - C(x, y)$ be the dual connective of a copula C. Then $H = \lambda C + (1 - \lambda)\overline{C}$ if and only if C satisfies, for all x, y in $[0, 1]$:*

$$C(x, y) = \frac{H(x, y) - (1 - \lambda)(x + y)}{2\lambda - 1}$$

Proof: $H = \lambda C + (1 - \lambda)\overline{C}$ is equivalent to $H(x, y) = \lambda C(x, y) + (1 - \lambda)(x + y - C(x, y))$, and this is equivalent to $C(x, y) = \dfrac{H(x, y) - (1 - \lambda)(x + y)}{2\lambda - 1}$.

Theorem 1 *Let H be a binary operation on $[0, 1]$ such that:*

- *H satisfies the following boundary conditions:*
  1. *$H(0, 0) = 0$, $H(1, 1) = 1$*
  2. *$H(x, 0) = H(0, x)$ for all $x \in [0, 1]$*
  3. *H has a linear behaviour on the boundaries*
- *H is 2-increasing if $\lambda > 1/2$ and 2-decreasing if $\lambda < 1/2$, where λ is the slope of the straight line $y = H(x, 1)$*

Then C defined by $C(x, y) = \dfrac{H(x, y) - (1 - \lambda)(x + y)}{2\lambda - 1}$ is a copula, and H can be written as $H = \lambda C + (1 - \lambda)\overline{C}$, where $\overline{C}$ is the dual copula.

Reciprocally, the connective $H = \lambda C + (1 - \lambda)\overline{C}$, with $\lambda \neq 1/2$, where C is a copula and $\overline{C}$ its dual copula, satisfies the two conditions of the former hypothesis.

Proof: Let $H(x, 0) = H(0, x) = ax + b$. The condition (1) implies $H(x, 0) = H(0, x) = ax$.

Let $H(x, 1) = mx + n$, $H(1, x) = px + q$. The condition (1) implies $H(x, 1) = mx + (1 - m)$ and $H(1, x) = px + (1 - p)$. Hence, $H(1, 0) = a$ and $H(1, 0) = 1 - p$, then $a = 1 - p$.

Analogously, $H(0, 1) = a$ and $H(0, 1) = 1 - m$, then $a = 1 - m$, and so $p = m$. Therefore, taking $a = 1 - \lambda$, the expression of H on the boundaries is:

- $H(x, 0) = H(0, x) = (1 - \lambda)x$
- $H(x, 1) = H(1, x) = \lambda x + (1 - \lambda)$

Let us now check that $C(x, y) = \dfrac{H(x, y) - (1 - \lambda)(x + y)}{2\lambda - 1}$ is a 2-copula:

$$C(x, 0) = \frac{H(x, 0) - (1 - \lambda)x}{2\lambda - 1} = \frac{(1 - \lambda)x - (1 - \lambda)x}{2\lambda - 1} = 0$$

$$C(x, 1) = \frac{H(x, 1) - (1 - \lambda)(x + 1)}{2\lambda - 1} = \frac{1 - \lambda + \lambda x - (1 - \lambda)x - (1 - \lambda)}{2\lambda - 1} = x$$

Analogously, $C(0, x) = 0$ and $C(1, x) = x$.

It remains to prove that C is 2-increasing. Let $x_1 < x_2$ and $y_1 < y_2$ be points of $[0, 1]$. C must satisfy the condition $C(x_2, y_1) - C(x_1, y_1) \leq C(x_2, y_2) + C(x_1, y_2)$, that is to say,

$$\frac{H(x_2, y_1) - (1 - \lambda)(x_2 + y_1) - H(x_1, y_1) + (1 - \lambda)(x_1 + y_1)}{2\lambda - 1} \leq \frac{H(x_2, y_2) - (1 - \lambda)(x_2 + y_2) - H(x_1, y_2) + (1 - \lambda)(x_1 + y_2)}{2\lambda - 1}$$

In the case $\lambda > 1/2$, this is equivalent to:

$$H(x_2, y_1) - (1 - \lambda)(x_2 + y_1) - H(x_1, y_1) + (1 - \lambda)(x_1 + y_1) \leq$$
$$H(x_2, y_2) - (1 - \lambda)(x_2 + y_2) - H(x_1, y_2) + (1 - \lambda)(x_1 + y_2),$$

or, what amounts to the same, to:

$$H(x_2, y_1) - H(x_1, y_1) \leq H(x_2, y_2) - H(x_1, y_2),$$

which is the condition of 2-increasing for H, which is true by hypothesis.

In the case $\lambda < 1/2$, it appears that H is 2-decreasing.

Finally, the expression $H = \lambda C + (1 - \lambda)\overline{C}$ is a direct result of the lemma.

Reciprocally, it is straightforward to check that the connective $H = \lambda C + (1 - \lambda)\overline{C}$, with $\lambda \neq 1/2$, C a copula and $\overline{C}$ its dual copula, satisfies the two conditions of the former hypothesis. □

When $\lambda \neq 1/2$, lemma 1 guarantees the uniqueness of C for a hybrid connective H.

When $\lambda = 1/2$ the hybrid combination simplifies in $H(x, y) = \frac{1}{2}(x + y)$. This is a special case, as we do not have the uniqueness of C. Indeed, many copulas, when combined using $\lambda = 1/2$, give $H = \frac{1}{2}(x + y)$. For example, with $C(x, y) = \min(x, y)$ and $C(x, y) = xy$.

The first condition of the theorem is just the translation of the boundary conditions due to the hybrid combination. The second condition shows that H increases (or decreases) like a 2-joint probability distribution function. This similarity with pdf's is because of the close connection between copulas and pdf's [3].

4 FUTURE WORK

From a theoretical point of view, future research is focused on two lines. On the one hand, on the study of a characterization in the case of a linear combination of a t-norm and its dual. On the other hand, taking into account the crucial role played by copulas in probability theory when computing multivariate distribution functions from margins, its seems reasonable to conjecture a theorem of existence of hybrid connectives defined starting from a copula and its dual to obtain multivariate possibility functions from margins in fuzzy sets theory.

ACKNOWLEDGEMENTS

This work has been partly funded by MCyT (Spanish Ministry of Science and Technology) MERITO project (TIC2002-04371-C02/01).

REFERENCES

[1] N. Agell, J.C. Aguado, *A hybrid qualitative-quantitative classification technique applied to aid marketing decisions.* Proc. ESOMAR Congress and Trade Exhibition 2001 Rome (Italy), (2001).

[2] J. Aguilar-Martin, N. Agell, M. Sánchez, F. Prats *Analysis of tensions in a Population Based on the Adequacy Concept.* Proc. 5th Catalonian conference on AI, Castellón, Spain, October 2002. Topics in Artificial Intelligence. LNAI 2504. Springer-Verlag (2002).

[3] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces.* North Holland, New York, (1983).

[4] J. Waissman, J. Aguilar, B. Dahhou and G. Roux, *Généralisation du degré d'adequation marginale dans la méthode de classification LAMDA.* 6èmes rencontres de la Société Francophone de Classification. Montpellier, France, (1998).