A note on measures of specificity for fuzzy sets

Didier Dubois   Henri Prade
Université Paul Sabatier
Laboratoire
Langages et Systèmes Informatiques
118, route de Narbonne
31062 - TOULOUSE CEDEX

Abstract

This note discusses a few issues related to Yager's specificity index and the "possibilistic entropy" recently introduced by Higashi and Klir. A probabilistic interpretation of Yager's index is provided ; moreover it is indicated that both indices can be straightforwardly extended to assess the amount of imprecision of Shafer's belief functions.

Key words : fuzzy set ; possibility distribution ; belief function ; imprecision ; entropy.

Following Zadeh [7] the contents of a proposition of the form 'X is A', which restricts (in a fuzzy or a non-fuzzy way) the possible values of a variable X on a universe of discourse S, is represented by means of the membership function of the subset induced by the predicate A ; this subset of S will be also denoted by A and its membership function by μ_A. Thus the proposition 'X is A' is translated into

$$\forall\, s \in S,\ \pi_X(s) = \mu_A(s) \qquad (1)$$

where π_X is the possibility distribution attached to the variable X.

Fuzzy set inclusion, defined by,

$$A \subseteq B \Leftrightarrow \forall s \in S,\ \mu_A(s) \leq \mu_B(s) \tag{2}$$

is in agreement with the fact that the larger A, the less restrictive the proposition p = 'X is A' (short for "the true value of X is in A") is for the possible values of X.

Yager [5], [6] introduced a so-called measure of specificity which estimates how precise is the information 'X is A' rather that its fuzziness (measured by the so-called entropy of the fuzzy set A, see [1]). A specificity measure Sp is such that, A and B being normalized fuzzy sets (i.e. $\exists\ s_1, s_2, \mu_A(s_1) = 1, \mu_B(s_2) = 1$)

$$\begin{cases} \text{i) } \forall A \subseteq S,\ Sp(A) \in [0,1] \\ \text{ii) } Sp(A) = 1 \Leftrightarrow A \text{ is a singleton of } S \\ \text{iii) } \forall A \subseteq B \Rightarrow Sp(A) \geq Sp(B) \end{cases} \tag{3}$$

When S is finite, Yager has proposed the following expression for defining the specificity

$$Sp(A) = \int_0^{\overline{\alpha}} \frac{1}{|A_\alpha|}\, d\alpha \tag{4}$$

where $\overline{\alpha} = \max_s \mu_A(s)$, $A_\alpha = \{s \in S, \mu_A(s) \geq \alpha\}$ and $|\ |$ denotes the cardinality. It can be easily checked that the expression defined by (4) satisfies the requirements (3). A crisp set can be less specific (i.e. precise) than a fuzzy set for restricting the possible values of a variable ; the maximum of specificity corresponds to a precise assessment of the value of the variable.

A measure of specificity must not be confused with a measure of fuzziness which estimates to what extent a subset has an ill-defined boundary. A measure of fuzziness is zero for crisp subsets and is maximum for the fuzzy set A defined by, $\forall s \in S,\ \mu_A(s) = \frac{1}{2}$. See [1].

In Dubois Prade [2], a one-to-one correspondence between a probability distribution p_X and a possibility distribution π_X is introduced on finite domains ; this transformation preserves the following inequalities

$$\forall F \in 2^S,\ N_X(F) \leq P_X(F) \leq \Pi_X(F) \qquad (5)$$

where P_X, Π_X and N_X are the probability, possibility and necessity measures based on p_X and π_X respectively (i.e. $P_X(F) = \sum_{s \in F} p_X(s)$; $\Pi_X(F) = \max_{s \in F} \pi_X(s)$; $N_X(F) = \min_{s \notin F} (1 - \pi_X(s))$. The inequalities (5) are in agreement with the motto : what is probable must be possible and what is necessary (i.e. ineluctable) must be probable.

The transformation is defined by

$$\forall S \in S,\ \pi_X(s) = \sum_{s' \in S} \min(p_X(s'), p_X(s)) \qquad (6)$$

and conversely $\forall\ i = 1,n,\ p_X(s_i) = \sum_{j=i}^{n} \frac{1}{j} \cdot (\pi_X(s_j) - \pi_X(s_{j+1}))$ (7)

where the n elements of S are supposed to be ordered according to the decreasing values of π_X and $\pi_X(s_{n+1}) = 0$ by convention. Note that

$$\forall\ i = 1,n,\quad p_X(s_i) \geq p_X(s_{i+1}) \Leftrightarrow \pi_X(s_i) \geq \pi_X(s_{i+1}) \qquad (8)$$

i.e. the transformation preserves the ordering among the elements of S.

Thus, a possibilistic interpretation can be provided for frequency histograms concurrently with the usual probabilistic interpretation.

The specificity of a normalized fuzzy set A, defined by (4) on a finite domain S, is still equal to

$$Sp(A) = \sum_{i=1}^{n} \frac{\mu_A(s_i) - \mu_A(s_{i+1})}{i} \qquad (9)$$

where the n elements of S are ordered according to the decreasing values of μ_A and $\mu_A(s_{n+1}) = 0$; thus, Sp(A) can be viewed as the probability of the element which has the greatest membership degree, in the sense of the transformation defined by (6) and (7) ; in other words, Sp(A) is the probability, computed from π_X using (7), of the most possible value of X when we interpret μ_A as π_X.

A possibility measure [7] based on a possibility distribution $\pi_X = \mu_A$ from S to [0,1] is a particular case of plausibility functions (see Shafer [4]). A plausibility function Pl is a set function which can be defined from a so-called basic probability assignment m, such that

$$\text{i) } m(\emptyset) = 0 \text{ ; ii) } \sum_{F \subseteq S} m(F) = 1 \qquad (10)$$

Then,

$$\forall G \subseteq S, \ Pl(G) = \sum_{G \cap F \neq \emptyset} m(F). \qquad (11)$$

By duality the belief function based on m is defined by $Bel(G) = 1 - Pl(\bar{G})$ where $\bar{G}$ is the opposite event of G.

When the subsets F of S which are such that m(F) >0 (and which are called "focal elements") are nested, the plausibility function is a possibility measure and then we have the following relation between the underlying possibility distribution and basic probability assignment (see [2])

$$\begin{cases} m_A(\{s_1, \ldots, s_i\}) = \mu_A(s_i) - \mu_A(s_{i+1}), \ i = 1,n \\ m_A(F) = 0 \text{ if } F \neq \{s_1, \ldots, s_i\} \end{cases} \qquad (12)$$

where the n elements of S are supposed to be ordered according to the decreasing values of μ_A ; $\mu_A(s_{n+1}) = 0$ by convention.

Note that formula (7) can be extended to provide a probabilistic approximation of any belief or plausibility function, under the form (see [2])

$$\forall s_i \in S, \ p_X(s_i) = \sum_{s_i \in F} \frac{m(F)}{|F|} \qquad (13)$$

The specificity of A defined by (4) is still equal to

$$Sp(A) = \sum_{F \subseteq S} \frac{m_A(F)}{|F|} \tag{14}$$

Thus, (14) still makes sense for general plausibility functions and not only for possibility measures.

Recently, Higashi and Klir [3] have introduced a possibilistic measure of uncertainty U defined by

$$U(A) = \int_0^1 \log_2(|A_\alpha|)\,d\alpha \tag{15}$$

where A is a normalized fuzzy set of S. When S is finite, using (12), (15) can be rewritten as

$$U(A) = \sum_{F \subseteq S} m_A(F) \,.\, \log_2(|F|) \tag{16}$$

The expression (16) shows that U(A) can be viewed as an Hartley measure of information weighted by the basic probability assignment m_A, since $\log_2(|F|)$ is the Hartley entropy of the crisp set F, see [3]. This is coherent with the interpretation (provided by (12)) of a fuzzy set as a weighted collection of nested crisp sets.

It is clear that (16) can be readily extended to a general plausibility function defined by its basic probability assignment m, as

$$U(m) = \sum_{F \subseteq S} m(F) \,.\, \log_2(|F|) \tag{17}$$

Probability measures are particular cases of plausibility functions whose focal elements reduce to singletons, see [4]. Thus U(m) is always zero in case of a probability measure since $m(F) > 0 \Rightarrow |F| = 1$. This remark and the interpretation of U(m) as a weighted Hartley entropy, indicate that U(m) estimates the imprecision of the focal elements and not at all to what extent these focal elements are conflicting ; the case of a probability measure corresponds to the maximum of precision of the focal elements and to the maximum of conflict between them, since the focal elements are singletons which do not overlap by nature ; in case of a possibility measure there is no conflict since the focal elements are nested and only the imprecision remains. Thus (16) or (17) clearly depart from Shannon entropy which, in some sense, assesses the amount of conflict of the available evidence regarding the values which some variable may take.

Indeed, it can be shown that U(A), defined by (16) satisfy the following requirements

i) $U(A) \in [0, +\infty)$

ii) $U(A) = 0 \Leftrightarrow$ A is singleton of S

iii) $A \subseteq B \Rightarrow U(A) \leq U(B)$ (see [3])

Thus, f(U(A)) where is a strictly decreasing one-to-one mapping form $[0, +\infty)$ to [0,1], can be used as a specificity measure concurrently with Sp(A) defined by (4).

Lastly, we may think of computing the Shannon entropy of the probability distribution issued by (17) from a possibility distribution. However, the measure of uncertainty which is thus attached to a possibility distribution is not monotonic with respect to the possibility distribution (i.e. fuzzy set) inclusion defined by (2), as it can be checked on counter-examples. Thus this approach cannot give birth to a measure of specificity in the sense of (3).

REFERENCES :

[1] - De Luca, A., Termini, S. (1972). A definition of a non-probabilistic entropy in the setting of fuzzy sets theory. Information and Control, Vol. 20, pp. 301-312.

[2] - Dubois, D., Prade, H. (1983) Unfair coins and necessity measures. A possibilistic interpretation of histograms. Fuzzy Sets and Systems, Vol. 10, n° 1, pp. 15-20.

[3] - Higashi, M., Klir, G.J. (1983) Measures of uncertainty and information based on possibility distributions. Int. J. General Systems, Vol. 9, n° 1, pp. 43-58.

[4] - Shafer, G. (1976) A Mathematical Theory of Evidence. Princeton University Press.

[5] - Yager, R.R. (1981) Measurement of properties on fuzzy sets and possibility distribution. Proc. 3rd Int. Seminar on Fuzzy Set Theory (E.P. Klement, Ed.), J. Kepler Univ., Linz, Austria, pp. 211-222.

[6] - Yager, R.R. (1983) An introduction to applications of possibility theory. Human Systems Management, Vol. 3, pp. 246-253.

[7] - Zadeh, L.A. (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets & Systems, Vol. 1, n° 1, pp. 3-28.