An Alternative Way of Calculating Uncertainty

Iliyana Raeva, Roumen Raev

Резюме: *В доклада са разгледани основните концепции, на базата, на които са изградени някои от съществуващите методи за измерване на размита (неточна) информация. Направен е кратък обзор на класическите мерки за неточна информация на Hartley и Shannon. Разгледани са два примера за обобщаване на тази теория и представяне на алтернативен метод за измерване на неопределена информация.*

Ключови думи: *размити множества, неточна информация, мерки за информация*

INTRODUCTION

The question of how measure vagueness or fuzziness has been one of the issues associate with the development of the theory of fuzzy sets. In general, a measure of fuzziness is a function **f : P(X) → R w**here **P(X)** denotes the set of all fuzzy subset A of X. That is, the function f assigns a value f(A) to each fuzzy subset A .

In order to quality as a meaningful measure of fuzziness, f must satisfy certain axiomatic requirements. Although not necessary unique, these requirements must fully capture the meaning of an intuitively acceptable characterization of the concept degree of fuzziness [1].

MEASURE OF FUZZINESS

Several measure of fuzziness have been proposed in the literature. [1]. One of them, perhaps the best known, is based on the following concepts:

1. The sharpness relation A< B is defined by:

$$\mu_A(x) \le \mu_B(x) \quad \text{for} \quad \mu_B(x) \le \frac{1}{2} \tag{1}$$

and

$$\mu_A(x) \ge \mu_B(x) \quad \text{for} \quad \mu_B(x) \ge \frac{1}{2} \quad \text{for all } x \in X. \tag{2}$$

2. The term maximally fuzzy defined by the membership grade ½ for all x∈ X.

This measure of fuzziness is defined by the function:

$$f(A) = -\sum_{x\in X}\left(\mu_A(x)\log_2 \mu_A(x)\right) + [1-\mu_A(x)]\log_2[1-\mu_A(x)]) \tag{3}$$

Its normalized version, f(A), for which 0 ≤ f(A) ≤1

Is clearly given by f(A) = f(A)/$|X|$ where $|X|$ denotes the cardinality of the universal set X.

Another measure of fuzziness, referred to as an index of fuzziness, is defined in terms of a metric distance (Hamming or Euclidean) of A from any the nearest crisp sets, say crisp set C, for which $\mu_C(x) = 0$ if $\mu_A(x) \le 1/2$ and $\mu_C(x) = 1$ if $\mu_A(x) > 1/2$.

When the Hamming distance is used, the measure of fuzziness is expected by the function:

$$f(A) = \sum_{x\in X}|\mu_A(x) - \mu_C(x)| \quad \text{for Euclidean distance:} \quad f(A) = \left(\sum_{x\in X}\left[\mu_A(x) - \mu_C(x)\right]^2\right)^{1/2} \tag{4}$$

REVIEW OF THE FUNDAMENTAL PROPERTIES OF THE CLASSIICAL MEASURA OF UNCERTAINTY

Two principle measure of uncertainty are recognized to the theory of fuzzy sets[3] One of them proposed by Hartley is based on the classical set theory. The other introduced by Shannon is formulated in terms of probability. Both of these measures pertain to some aspects of ambiguity, as opposed to vagueness or fuzziness. Harley's measure pertains to nonspecificity, Shannon's measure to conflict or dissonance in evidence.

Both Harley and Shannon introduced their measure for the purpose of meaning information in term of uncertainty. These measures are often referred to as measures of information.

Hartley information can also characterized by the following axioms [4]:

A1 (additively) I(M.N) = I(M)+I(N) for all M and M € Ž
A2 (monotonic) I(M) ≤ I(M+1) for all M € Ž
A3 (normalization) I(2) = 1

As expressed by the following uniqueness theorem, the Hartley information is the following theorem:

Theorem1 Function I(M) = $\log_2 N$ is the only function that satisfied Axiom 1-3.

The ***Shannon*** entropy[3], which is a measure of uncertainty and information formulate in terms of probability theory, is expressed by the function

H(p(x)/x€ X) = - $\sum_{x\in X} p(x).\log_2 p(x)$, where p(x)/ x€ X is a probability distribution on a finite set X. It is thus a function of the form H: R→ [0, ∞) where R denotes the set of all probability distribution on finite sets.

AN ALTERNATIVE WAY OF CALCULATING UNCERTAINTY

1. Summary of the Hartley's method

Information translation can be generalizes to express the constrain among more than two sets. It is always expressed as the difference between the total information based on the individual sets and the information. Formally

$$T(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} I(X_i) - I(X_1, X_2, \ldots, X_n) \qquad (5)$$

We will illustrate the method by example:

Example 1

Consider two variables x and y whose value are taken from sets X= {low, medium, high} and Y= {1, 2, 3, 4}, respectively. It is known that the variables are constrained by the relation R expressed by the matrix medium $\begin{array}{c} Low \\ Medium \\ Hogh \end{array} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$. We can set that the low value of x does not constrain y at all, the medium value of x constrains y partially, and the high value constrains it totally. The following types can be calculates:

I(X) = $\log_2|X|$ = $\log_2 3$=1.6

I(Y) = $\log_2|Y|$ = $\log_2 4$=2

I(X,Y) = $\log_2|R|$ = $\log_2 7$=2,8

I(X/Y) = I(X, Y) – I(Y) = 2.8 – 2 = 0.8
I(Y/X) = I(X, Y) – I(X) = 2.8 – 1.6 = 1.2

T(X,Y) =I(X)+I(Y) - I(X, Y) = 1.6 +2 -2.8 = 0.8

The information is based on uncertainty associated with a choice among a certain number of alternatives.

Example 2
Let the set $X = \{x_1, x_2, x_3, x_4\}$ with the probability distribution

$$p = \left(p_1 = .25,\ p_2 = .5,\ p_3 = .125,\ p_4 = .125 \right)$$

be given where p_i denotes the probability of x_i for all $i \in \mathrm{N}_4$. Consider the four branching schemes specified in IV for calculating the uncertainty of this probability distribution. Employing the branching property of Shannon entropy, the resulting uncertainty should be the same regardless of which of the branching schemes we use. Let us perform and compare the four schemes of calculating the uncertainty.

Scheme I.
According to this scheme, we calculate the uncertainty directly: $H(p) = -.25 \log_2 .25 - .5 \log_2 .5 - 2\ \mathrm{X}\ .125 \log_2 .125 = .5 + .5 + .375 + .375 = 1.75.$

Scheme II.

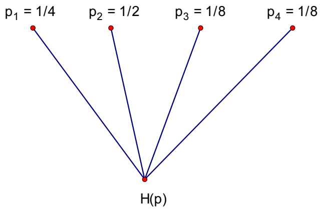
$$H(p) = H(p_A, p_B) + p_A H(p_1 / p_A, p_2 / p_A) + p_B H(p_3 / p_B, p_4 / p_B) =$$
$$H\left(\frac{3}{4}, \frac{1}{4}\right) + .75 H\left(\frac{1}{3}, \frac{2}{3}\right) + .25 H\left(\frac{1}{2}, \frac{1}{2}\right) = .811 + .689 + .25 = 1.75$$

Scheme III.

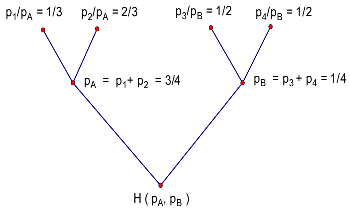
$$H(p) = H(p_1, p_A) + p_A H(p_2 / p_A,\ p_3 / p_A,\ p_4 / p_A) = H\left(\frac{1}{4}, \frac{3}{4}\right) + .75 H\left(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\right) = .811 + .939 = 1.75$$

Scheme IV.

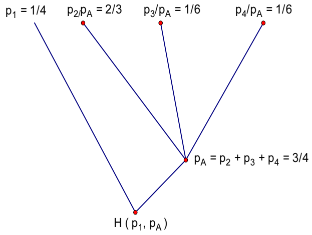
$$H(p) = H(p_1,\ p_A) + p_A H(p_2 /\ p_A,\ p_B /\ p_A) + p_B H(p_3 /\ p_B,\ p_4 /\ p_B) =$$
$$H\left(\frac{1}{4}, \frac{3}{4}\right) + .75 H\left(\frac{2}{3}, \frac{1}{3}\right) + .25 H\left(\frac{1}{2}, \frac{1}{2}\right) = .811 + .689 + .25 = 1.75.$$



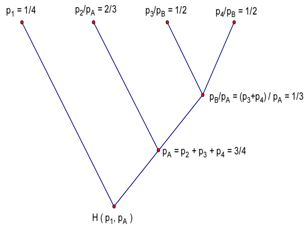
<u>**Scheme I**</u>



<u>**Scheme II**</u>



Scheme III



Scheme IV

These results thus demonstrate that the uncertainty can be calculated in terms of any branching scheme. There are, of course, many additional branching schemes in this example, each of which can be employed for calculating the uncertainty and each of which must lead to the same result.

LITERATURE

[1] Alsina, C., E. Trillas, and L. Valvedre [1983], “On some logical connectives for fuzzy set theory.” *J. of Math. Analysis and Applications*, 93

[2] Ashby, W.R. [1965], “Measuring the internal informational exchange in a system” *Cybernetic a,* 1

[3] Buoncristiani, J. F. [1983], “Probability on fuzzy sets.” *J. of Math. Analysis and Applications,* 96

[4] Jones, D.S. [1979], *Elementary Information Theory.* Clarendon Press, Oxford.

За контакти:

гл. ас. .Илияна Раева катедра “Алгебра и геометрия”, Русенски университет “Ангел Кънчев”, тел.: 082-888 453, e-mail: iraeva@yahoo.com

гл. ас. .Rumen Раев катедра “Математичен анализ”, Русенски университет “Ангел Кънчев”, тел.: 082-888 226, e-mail :rumraev@yahoo.com

Докладът е рецензиран.