

A (NEW) MEASURE OF FUZZY UNCERTAINTY VIA INTERVAL ANALYSIS, WHICH IS FULLY CONSISTENT WITH SHANNON THEORY

GUY JUMARIE

Abstract. Many authors have suggested different measures of the amount of uncertainty involved in fuzzy sets, but most of these concepts suffer from drawbacks: mainly, they are indexes of fuzziness rather than measures of uncertainty, and they are not fully consistent with Shannon theory. The question is herein once more considered by combining the information theory of deterministic functions, recently initiated by the author, with the viewpoint of interval analysis; and one so derive the new concept of "uncertainty of order c of fuzzy sets". It is shown that it satisfies the main properties which are desirable for a measure of uncertainty.

Some topics are outlined, such as informational distance between fuzzy sets, and mutual information between fuzzy sets for instance. One so has at hand a unified approach to Shannon information expressed in terms of probability, and to fuzzy information described by weighting coefficients commonly referred to as possibility distribution.

1. Introduction

The last few years have witnessed a number of critical papers on the goals, the foundations and the present development of the so-called fuzzy set theory [2], [7], [11], [12], [16], [21], [22], [24], [25], [26], [29], [33] which seem to indicate that the latter has not yet achieve its final form and is still looking for improvement in its interpretations and applications.

This theory is still suffering from a rather careless mixture of concepts and methods (subjective probability, objective probability, possibility, min-max operator, ...) and this feature is mainly transparent in the various models of entropy which have been proposed in order to measure the amount of uncertainty involved in fuzzy sets. Clearly, there is an obvious confusing between the uncertainty which is caused by randomness (if there

Received June 4, 1990

Research supported by the National Research Council of Canada

Keywords: Information, Uncertainty, Shannon entropy, Hartley entropy, Membership Functions, Possibility.

is any) on the one hand, and by fuzziness, that is to say vagueness of definition, on the other hand. Nevertheless, a careful derivation of a good measure of fuzzy set uncertainty is of interest because (i) it would enlarge the area of applications of the maximum entropy principle [13],[14]; (ii) it would provide new ways to meaningfully address identification problems; (iii) and more generally it would allow us to extend Shannon information theory to the fuzzy set framework in quite a consistent manner.

As it has been pointed out by several authors, the main source of the difficulties above mentioned is that the very nature of the so-called membership function of a fuzzy set is not yet quite clarified in the literature. Fuzziness appears only in linguistics in the broad sense of this term, and we have suggested [19] a modelling via coupling effects between syntax and semantics in the framework of a relative information theory. In this approach, a membership function cannot be identified with a probability distribution, or with subjective probability, and in this way we fully agree with fuzzy scientists.

As a matter of fact, it appears that some of the theoretical difficulties so encountered can be circumvented if we merely utilize interval analysis [23] and it is exactly the purpose of the following, to expand this idea.

The paper is organized as follows. After a critical review of some entropic concepts which have been proposed to deal with fuzzy sets, we shall give a few background on our recent information theory of deterministic functions. Then we shall apply it to fuzzy sets and we shall so obtain the concept of U -uncertainty of order c , and later we shall outline some consequences and applications.

2. A Critical Review of Fuzzy Entropic concepts

2.1 Preliminary Background

For a discrete random variable $X \in R$ which takes on the values (x_1, x_2, \dots, x_n) with the probabilities (p_1, p_2, \dots, p_n) , the Shannon entropy or Sentropy $H(X)$ is defined as

$$H(X) := - \sum_{i=1}^n p_i \ln p_i \quad (2.1)$$

and measures the amount of uncertainty (and incidentally of information, but incidentally only) involved in the probability distribution $\{p_i\}$ or equivalently in X .

The amount of uncertainty involved in an incomplete probability distribution, namely in $(p_i, p_{i+1}, \dots, p_{i+k})$ is

$$H(X; i, i+k) := - \left(\sum_{j=0}^k p_{i+j} \ln p_{i+j} / \sum_{j=0}^k p_{i+j} \right) \quad (2.2)$$

and this expression is consistent with the fact that the uncertainty involved in the event $X = x_i$ is $-\ln p_i$ (instead of $-p_i \ln p_i$). For further details, see for instance the Ref. [1].

For a continuous random variable $X \in R$ with the probability density $p(x)$, one has

$$H(X) := - \int_R p(x) \ln p(x) dx \tag{2.3}$$

and

$$H(X; \Omega) := - [\int_{\Omega} p(x) \ln p(x) dx / \int_{\Omega} p(x) dx] \tag{2.4}$$

Prerequisites for any generalization of Shannon's measure of uncertainty. Our claim is that, any extension of the Shannon entropy which aims to involve some external parameters which characterize uncertainty different from randomness, should comply with two requirements.

(i) First, of course, it must yield (2.1) and (2.3) as special cases, when the parameters are given some special actual physical values;

(ii) and second, it must be consistent with the fact that the uncertainty involved in one event only, say $X = x$, is $-\ln p(x)$.

2.2 From Fuzzy-Set Entropy to Fuzzy Set-Entropy

(i) The first formal extension of entropy has been proposed by Belis and Guiasu [3] in the form of the so-called useful entropy

$$H(X, V) := - \sum_{i=1}^n v_i p_i \ln p_i \tag{2.5}$$

where $V := (v_1, v_2, \dots, v_n)$, $v_i \geq 0$, is referred to as a utility distribution.

The trouble is that $H(X, V)$ so defined does not satisfy the requirement (ii) above. Indeed, if we make $v_k = 1$ and $v_i = 0$, $i \neq k$, into (2.5) we find that the uncertainty involved in the event $X = x_k$ is $-p_k \ln p_k$ instead of $-\ln p_k$ as it should be. As a matter of fact, the Shannon entropy is a weighted mean value of local event uncertainties, and with this feature in mind, it would be more convenient to take the alternative definition

$$\tilde{H}(X, V) := - \sum_{i=1}^n v_i p_i \ln p_i / \sum_{i=1}^n p_i v_i \tag{2.6}$$

(ii) Xadeh [32] introduced a concept of "entropy of fuzzy event" as follows. Let be given the fuzzy set defined by the membership function $f(x_i)$, $x_i \in \{x_i\}$, $i = 1, 2, \dots, n$, $0 \leq f(x_i) \leq 1$, on the one hand, and the probability distribution $p(x_i)$ on the other hand; then this entropy is defined by the expression

$$H(f, p) := \sum_{i=1}^n f(x_i) p(x_i) \ln p(x_i) \tag{2.7}$$

The surprising-looking consequence of this definition is that the uncertainty involved in a fuzzy set is lower than that of the corresponding crisp set! In other words, fuzzyfying

a crisp set diminishes the amount of uncertainty we may have about it! It is clear that this drawback would be less transparent if instead we had (for instance) the definition

$$\bar{H}(f, p) := - \sum_{i=1}^n f(x_i)p(x_i) \ln p(x_i) / \sum_{i=1}^n f(x_i)p(x_i) \quad (2.8)$$

(iii) De Luca and Termini [5] introduced a quantity $H(f)$ referred to as entropy of fuzzy sets, which is defined by the expression

$$H(f) := \sum_{i=1}^n h(f(x_i)) \quad (2.9)$$

where $h(f)$ denotes the entropy (first utilized by Fermi)

$$h(f) := -f \ln f - (1-f) \ln(1-f) \quad (2.10)$$

This definition gives rise to two disturbing questions. First, in view of (2.10), the pair $(f, 1-f)$ is merely considered as a probability distribution, which contradicts the well known claim of fuzzy scientists themselves. More important is the following second point. Assume that we refine the definition of the fuzzy set by measuring the value of $f(\xi)$ at a new point ξ , $x_i < \xi < x_{i+1}$, in such a manner that the set has the new membership function $\tilde{f} := \{f(x_1), \dots, f(x_i), f(\xi), f(x_{i+1}), \dots, f(x_n)\}$. We would then have $H(\tilde{f}) > H(f)$. In other words, we increase the amount of uncertainty we have about the fuzzy set just because we are refining its definition! It is clear that we would not have this trouble if $H(f)$ were an average of uncertainty. For instance, it would already be better to take

$$\bar{H}(f) := \frac{1}{n} \sum_{i=1}^n h(f(x_i)) \quad (2.11)$$

In fact, strictly speaking, $H(f)$ is not an informational entropy in Shannon sense, that is to say a measure of uncertainty, but rather an index of fuzziness which measures the discrepancy of the fuzzy set with respect to a crisp set.

(iv) Later, again De Luca and Termini [6] introduced the entropy

$$H_T(f, p) := H(f) - \sum_{i=1}^n p_i \ln p_i \quad (2.12)$$

which is an obvious adaptation of the well known information theoretic law in accordance of which one has $H(X, Y) = H(X) + H(Y)$ provided that X and Y are independent. But unfortunately, strictly speaking, this expression is not homogeneous in the physical sense of this term. Indeed, at first glance, one has

$$H_T(f, p) := H(f) + H(X),$$

but as we just mentioned it above, $H(f)$ should be considered as a divergence, in Kullback sense, rather than as an entropy.

(v) Hirota [10] suggested a mixed probabilistic-fuzzy entropy in the form

$$H_H(f, p) := - \sum_{i=1}^n p_i \ln p_i + \sum_{i=1}^n p_i h(f(x_i)) \tag{2.13}$$

which is an obvious adaptation of the well known information theoretic equation $H(X, Y) = H(X) + H(Y/X)$. But unfortunately, $H_H(f, p)$ combines all the flaws above mentioned. Mainly, the more we refine the definition of the fuzzy set, the more we have uncertainty about it!

(vi) In an attempt to unify the various entropic models which have been proposed in the literature, including the above ones, Bouchon [4] considers simultaneously two random variables X and Y and their respective probability distributions P and Q . X represents the posterior characterization of events, and Y describes a prior one. By using a set of given prior axioms which are chosen in an ad hoc manner, she can derive the entropies (2.7) and (2.9). But according to our previous remarks, it is clear that these axioms are not suitably defined since they lead to consequences which contradict which contradict basic results of Shannon theory!

(vii) Higashi and Klir [9] introduced an entropic model for fuzzy intervals as follows. Let us consider (just to simplify the writing) a continuous fuzzy set with the membership function $f(x)$, $0 \leq f(x) \leq 1$. Let us construct the interval $i(z) := \{x : f(x) \geq z\}$ where z denotes the level of set of f , and let $L(z)$ denote its length; then the entropy $U_0(f)$ of the fuzzy set f is defined by the expression

$$U_0(f) := \frac{1}{\bar{f}} \int_0^{\bar{f}} \ln L(z) dz \tag{2.14}$$

with $\bar{f} := \max f(x), x \in R$.

For a discrete fuzzy set, we have the expression

$$U_0(f) := \frac{1}{\bar{f}} \sum_{i=1}^{n-1} (z_{i+1} - z_i) \ln L(z_{i+1}) \tag{2.15}$$

which can be re-written in the form

$$U_0(f) = - \sum_i z_i \Delta \ln L(z_i) \tag{2.16}$$

with $\Delta \ln L(z_i) := \ln L(z_{i+1}) - \ln L(z_i)$.

This quantity is symmetric, expansible, continuous, additive for an operation and defined by the minimum of the values of the grades of memberships. Moreover, it is the

average value of the Hartley entropies $\ln L(z_i)$ in such a manner that it does not suffer from the drawback involved in the various measures of uncertainty above.

But despite its interesting features, this definition appeals to the following remark. In fact, the genuine meaning of $U_0(f)$ is that it measures an amount of information (or transinformation in Shannon's terminology) instead of an amount of uncertainty. Indeed, if we refer to equation (2.16), $U_0(f)$ appears to be the average value of differences of uncertainties, clearly $\Delta \ln L(z)$, that is to say the average value of measures of information! Another way to assess the matter is as follows. Assume that we rewrite $U_0(f)$ in the form

$$U_0(f) = \frac{1}{f} \sum_i z_i \ln[L(z_{i+1})/L(z_i)] \quad (2.17)$$

then it looks like a Kullback divergence.

(vii) Recently Ramer [27] suggested that a more general definition of "information function" for fuzzy sets would be in the form

$$I(f) := \sum_i [\tau(f_{i+1}) - \tau(f_i)] \ln x_{i+1} \quad (2.18)$$

where τ is a nondecreasing function of $[0, 1]$ onto itself. If one also requires that $I(f)$ satisfies some form of a linear property, like branching condition, one arrives at the particularly simple expression

$$I(f) = \sum_i (f_{i+1} - f_i) \ln x_{i+1} \quad (2.19)$$

$$= - \sum_i \Delta \ln x_{i+1} \quad (2.20)$$

We shall comment on this result as follows. In his derivation, the author considers $\{f_i\}$ as being a so-called possibility distribution, that is to say such that $f_1 \leq f_2 \leq \dots \leq f_n = 1$. But a special case of such possibility distribution is the cumulative distribution function

$$f_n := \sum_{i=1}^n p_i \quad (2.21)$$

where $\{p_i\}$ is a probability distribution. If then we substitute (2.21) into (2.19), we obtain

$$\begin{aligned} I(f) &= \sum_i p_{i+1} \ln x_{i+1} \\ &= E\{\ln X\}! \end{aligned}$$

We believe that all these remarks and comments are a sufficient motivation to once more re-consider the question of how to measure the amount of uncertainty involved in a fuzzy set, and we herein propose an alternative via information of deterministic functions.

3. Informational Entropy of Deterministic Functions

3.1 Main Definitions and Motivation

Entropy of mappings. Let $f : R^n \rightarrow R^n$, $x \rightarrow f(x)$ denote a continuously differentiable mapping, the Jacobian determinant of which is denoted by $f'(x)$; then a measure $H_c(f(\cdot); \Omega)$ of the amount of uncertainty it involves on the domain Ω , and which furthermore is fully consistent with Shannon entropy of random variables, is the S -entropy (S holds for Shannon) of order $c \in R$, defined by the expression

$$H_c(f(\cdot); \Omega) := \frac{\int_{\Omega} |f'(x)|^c \ln |f'(x)| dx}{\int_{\Omega} |f'(x)|^c dx} \tag{3.1}$$

Entropy of family of mappings. Let $f : \Omega \times D \rightarrow R^n$, $\Omega \subset R^n$, $D \subset R^m$, $(x, z) \rightarrow f(x, z)$ denote a function which is continuous with respect to z and continuously differentiable *w.r.t.* to x , and the Jacobian determinant of which is denoted by $f_x(x, z)$. Then as a direct consequence of the definition of Shannon entropy of random variables, its S -entropy of order $(b, c(z))$, $b \in R$, $c(z) \in R$ on the domain $\Omega \times D$ is given by the expression

$$H_{b,c(z)}(f_z(\cdot); \Omega \times D) = \frac{\int_D e^{bH_{c(z)}(f(\cdot, z); \Omega)} H_{c(z)}(f(\cdot, z); \Omega) dz}{\int_D e^{bH_{c(z)}(f(\cdot, z); \Omega)} dz} \tag{3.2}$$

On the derivation of these definitions. These measures of uncertainty are not “ad hoc” definitions made for convenience only, but rather, they are direct consequences of the expression of Shannon entropy of random variables, and as such, they are fully consistent with Shannon theory. Loosely speaking, the derivation of the first expression works as follows.

(i) For two random variables X and Y one has

$$H(X, Y) \leq H(X) + H(Y)$$

where the equality holds when and only when X and Y are independent. We may therefore think of $H(Y)$ as

$$H(Y) = \max_{p(x)} [h(X, Y) - H(X)] \tag{3.3}$$

where $p(x)$ is the probability density of X , subject to the condition

$$H(X) = h.$$

(ii) Next, assume now that one has $Y = f(X)$, then it is well known that

$$H(Y) = H(X) + \int_{R^n} p(x) \ln |f'(x)| dx \tag{3.4}$$

If we think of $H(Y)$ as

$$H(Y) =: H(X, f(.)) \quad (3.5)$$

then, according to (3.3) we shall set

$$\begin{aligned} H(f(.)) &:= \max_{p(x)} [H(Y) - H(X)] \\ &= \max_{p(x)} \int_{R^n} p(x) \ln |f'(x)| dx \end{aligned} \quad (3.6)$$

therefore the expression (3.1).

The derivation of (3.2) is based on a randomization of the parameter z . For more details, see the Ref. [18].

3.2 Some Useful Remarks

(i) The constant c in (3.1) and the pair $(b, c(z))$ in (3.2) are related to the Lagrange parameters involved in the maximization expressed by (3.6).

(ii) Assume that $X \in R$, and that $f(x)$ is the cumulative distribution function $F(x)$ of X ; i.e. $F'(x) = p(x)$, then one has

$$H_0(F(.); -L, +L) = \frac{1}{2L} \int_{-L}^{+L} \ln p(x) dx \quad (3.7)$$

and

$$H_1(F(.); R) = -H(X) \quad (3.8)$$

This remark underlines two points of importance. First, the most natural value, which is consistent with Shannon theory, for c in (3.1) and $(b, c(z))$ in (3.2) is

$$b = c = 1, \quad c(z) \equiv 1;$$

and second, the corresponding entropies (3.1) and (3.2) then clearly exhibit the definition of the weighting function involved in the averaging of the entropic density $\ln|f'(x)|$.

(iii) The meaning of (3.1) should be understood as follows. If the basic object we are observing is the whole pattern $\{f(x); x \in \Omega\}$, then its mean uncertainty is $H_c(f(.); \Omega)$. But if we assume that the basic object of the observation process is the point $(x, f(x))$, then the total amount of uncertainty involved in the set $\{f(x); x \in \Omega\}$ is $\int_{\Omega} |f'| \ln|f'| dx$. For a random variable, these two measures of uncertainty are equal since then one has $\int_R p(x) dx = 1$.

4. U_c -Uncertainty of Fuzzy Sets

4.1 Preliminary Definitions and Notations

In R , we consider a fuzzy interval defined by the membership function $f(x)$, $0 \leq f(x) \leq 1$, $x \in R$. We shall assume that the fuzzy set is bounded (or finite), and that it is convex, that is to say $f(x)$ is convex.

For any level of set z , $0 \leq z \leq 1$, the equation $z = f(x)$ has two solutions $x_m(z)$ and $x_M(z)$, $z = f(x_m) = f(x_M)$, such that $x_m \leq x_M$. We define the interval length

$$L(z) := x_M(z) - x_m(z).$$

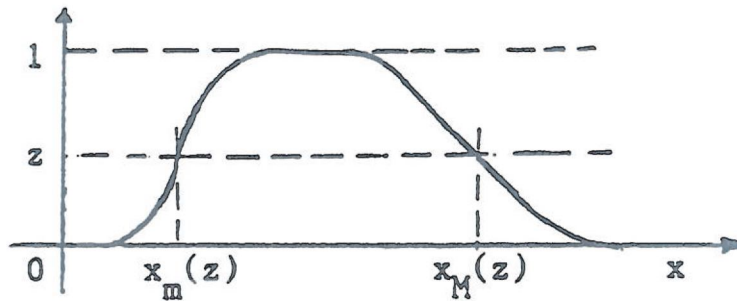


Fig 1. Fuzzy set and membership function

The significance of $f(x)$ in practical problems is not yet quite clear, fuzzy scientists refer to so-called possibility distributions, and we shall herein merely consider the fuzzy set $f(x)$ as a family of intervals $[x_m(z), x_M(z)]$ indexed by the parameter z . $f(z)$ is then simply a weighting coefficient which is ascribed to $L(z)$.

4.2 U_c -Uncertainty of Continuous Fuzzy Sets

Main Definition. Let $f(x)$ denote a convex finite fuzzy interval with $\max f(x) = 1$, $x \in R$. Then as a direct consequence of the definition of Shannon entropy of random variables, a measure of the amount of uncertainty it involves is the *S-uncertainty of order c* , $c \in R$ defined by the expression

$$U_c(f) := \frac{\int_0^1 L^c(z) \ln L(z) dz}{\int_0^1 L^c(z) dz} \tag{4.1}$$

Derivation of this definition. (i) We are dealing with a family of crisp sets indexed by z , and we thus refer to the equation (3.2).

(ii) We identify the uncertainty involved in a bounded crisp set $[a, b]$ with the uncertainty of the uniform probability density $1/(b - a)$ to obtain $\ln(b - a)$.

(iii) We can now apply equation (3.2) in which we set $H_c(f(\cdot, z); \Omega) \equiv \ln L(z)$.

4.3 U_c -Uncertainty of Discrete Fuzzy Sets

Assume that we are now dealing with a discrete convex fuzzy set $\{f(x_1), f(x_2), \dots, f(x_n)\}$, and let $\{z_1 = 0, z_2, \dots, z_n = 1\}$ denote the levels of set of f . Then as a consequence of the information theory of deterministic functions, the discrete counterpart of the distributed entropy (3.2) directly yields the S -uncertainty of $\{f(x_i)\}$ in the form

$$U_c(f) := \frac{\sum_{i=1}^n L^c(z_i) \ln L(z_i)}{\sum_{i=1}^n L^c(z_i)} \quad (4.2)$$

Remark that there is a complete consistency between (4.1) and (4.2). Indeed, if we assume that $z_i - z_{i-1} = \text{constant}$, then on multiplying (4.2) by the term $(z_i - z_{i-1}) / (z_i - z_{i-1})$ we then have

$$\lim U_c(f(\cdot))_{(4.2)} = U_c(f(\cdot))_{(4.1)} \text{ as } z_i - z_{i-1} \downarrow 0.$$

Comments. Assume that $c = 0$ in equation (4.1), then we obtain a measure of uncertainty which is equivalent to the uncertainty $U_0(f)$ expressed by (2.14). If we make $c = 0$ in (4.2), we get

$$U_0(f) = \frac{1}{n} \sum_{i=1}^n \ln L(z_i) \quad (4.3)$$

and this expression is equal to (2.15) in the special case when $z_{i+1} - z_i$ is constant.

5. Some Properties of U_c -Uncertainty

5.1 Preliminary Remarks

Although our derivation of U_c -uncertainty of fuzzy intervals is rigorously supported by Shannon information theoretic framework via entropy of deterministic functions, we still need to verify that it satisfies suitable likely properties for a satisfactory measure of uncertainty. But on doing so, care must be exercised, and we should have in mind that $U_c(f)$ so defined is the uncertainty involved in the basic object we are dealing with (indeed the fuzzy set is herein considered as the first element of the theory) and should not be thought of as the result of any composition law describing how local uncertainties are aggregated.

In this way of thought, we should rather refer to the basic suitable properties that *Shannon himself* suggested for the entropy of random variables, and which we herein bear in mind for convenience.

Consider a discrete random variable X with the probability distribution (p_1, p_2, \dots, p_n) , and let $\Phi(p_1, p_2, \dots, p_n)$ denote a measure of the amount of uncertainty it involves. Then $\Phi(\cdot)$ should satisfy the following properties.

(A1): $\Phi(\cdot)$ should be a continuous function of its arguments.

(A2): When all the probabilities are equal, $p_1 = p_2 = \dots = p_n = 1/n$, then $\Phi(\cdot)$ should be an increasing function of n .

(A3): Assume that a given outcome can be considered as a sequence of two events; then the whole uncertainty $\Phi(\cdot)$ should be the weighted combination of the partial uncertainties involved in $(p_1, p_2, \dots, p_{n-1} + p_n)$ and $[p_{n-1}/(p_{n-1} + p_n), p_n/(p_{n-1} + p_n)]$ respectively.

For instance, for three events, one should have the relation

$$\Phi(p_1, p_2, p_3) = \Phi(p_1, p_2 + p_3) + (p_2 + p_3)\Phi\left(\frac{p_2}{p_2 + p_3}, \frac{p_3}{p_2 + p_3}\right) \quad (5.1)$$

In the sequel, when we shall enumerate the properties of $U_c(f)$, we should have in mind to compare some of them with (A1)-(A3).

5.2 Some Properties of $U_c(f)$

(P1) $U_c(f)$ is a continuous function of $L_i := L(z_i)$ for every i . (This property compares with (A1)).

(P2) Assume that $L(z) = L$ is constant, in other words that we are dealing with a crisp set; then one has

$$U_c(f) = \ln L$$

and $U_c(f)$ is an increasing function of L . (This property compares with (A2)).

(P3) For a discrete fuzzy set, let us define (see equation (4.2))

$$\begin{aligned} U_1(f) &=: U_S(f), \text{ } S \text{ holds for Shannon} \\ &=: \Psi(L_1, L_2, \dots, L_n) \end{aligned} \quad (5.2)$$

and

$$L := \sum_{i=1}^n L_i \quad (5.3)$$

then the following relation holds,

$$\Psi(L_1, L_2, \dots, L_n) = \Psi(L_1 + L_2, L_3, \dots, L_n) + \frac{L_1 + L_2}{L} \Psi\left(\frac{L_1}{L_1 + L_2}, \frac{L_2}{L_1 + L_2}\right) \quad (5.4)$$

This property compares with (A3) above, and its meaning can be exhibited as follows. Assume that we randomize the set of the n intervals in (4.2) by using the probability distribution $(L_1/L, L_2/L, \dots, L_n/L)$, then the corresponding Shannon entropy $\Phi(\cdot)$ is such that

$$\Phi(L_1/L, L_2/L, \dots, L_n/L) + \Psi(L_1, L_2, \dots, L_n) = \ln L \quad (5.5)$$

and as a result, equation (5.1) direct yields (5.4).

Assume that L is constant; then Φ is maximum when $L_1 = L_2 = \dots = L_n = L/n$ to yield

$$\Phi(L/n, L/n, \dots, L/n) = \ln n \quad (5.6)$$

and correspondingly

$$\Psi(L/n, L/n, \dots, L/n) = \ln(L/n) \quad (5.7)$$

These relations illustrate the relative significance of the U_c -uncertainty with respect to the usually H entropy.

(P4) Let $f(x)$ and $g(x)$ denote two fuzzy sets such that

$$L_g(z) = e^k L_f(z) \quad (5.8)$$

where k denotes a constant, then one has

$$U_c(g) = U_c(f) + k \quad (5.9)$$

(P5) Let \bar{L} and \underline{L} be defined as

$$\bar{L} := \max L(z), \quad z \in [0, 1]$$

$$\underline{L} := \min L(z), \quad z \in [0, 1]$$

then $U_c(f)$ is an increasing function of c and one has

$$\lim U_c(f) = \ln \bar{L} \quad \text{as } c \uparrow +\infty \quad (5.10)$$

$$= \ln \underline{L} \quad \text{as } c \downarrow -\infty \quad (5.11)$$

The proof of this result is based on the following remark. Consider the integral

$$J = \int_a^b \alpha^c(x) \beta(x) dx, \quad \alpha > 0$$

then when c increases, the main contribution to the value of J is provided by $\alpha(\bar{x})$ where

$$\alpha(\bar{x}) := \max \alpha(x), \quad x \in [a, b]$$

in such a manner that, loosely speaking, for large values of c , one has

$$J = \int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon} \alpha^c(\bar{x}) \beta(x) dx$$

where ε denotes a small positive constant.

6. Additivity and Sub-Additivity of U_c -Uncertainty

6.1 On the Min-Max Operator

Let A and B denote two fuzzy sets defined by their respective membership functions $a(x)$ and $b(x)$, $x \in R$. It is by-now classical in fuzzy set theory to take the definition

$$A \cup B := \max\{(a(x), b(x)), x \in R\} \quad (6.1)$$

and

$$A \cap B := \min\{(a(x), b(x)), x \in R\} \quad (6.2)$$

and most authors who dealt with fuzzy uncertainty, claimed that a good measure of such uncertainty should be consistent with these min-max operators, and as a result, they selected prior axioms in this way. For instance, they asked for additivity and sub-additivity as defined below.

Additivity. Given two domains X and Y , and two independent membership functions $f : X \rightarrow [0, 1]$, $g : Y \rightarrow [0, 1]$, define the joint membership function $f \otimes g : (x, y) \rightarrow \min(f(x), g(y))$, then one should have

$$U(f \otimes g) = U(f) + U(g) \quad (6.3)$$

Sub-additivity. Given two domains X and Y and a membership function $f : X \times Y \rightarrow [0, 1]$ on the product space, define the marginal membership functions f_Y on X and f_X on Y as

$$f_Y(x) := \max_{y \in Y} f(x, y), \quad y \in Y \quad (6.4)$$

$$f_X(y) := \max_{x \in X} f(x, y), \quad x \in X \quad (6.5)$$

then one should have

$$U(f) \leq U(f_X) + U(f_Y) \quad (6.6)$$

The trouble is that the min-max operators (6.1) and (6.2) also have been and still are objects of contention [30]. In some special frameworks, it has been shown that these definitions of fuzzy union and fuzzy intersection are the only natural and reasonable definitions extending the standard set theory notion of union and intersection [31], but unfortunately, the same type of proof that demonstrates this property, has failed to demonstrate the corresponding result for the extension of complement as defined by fuzzy scientists, clearly

$$\overline{A} := \{(1 - a(x)), x \in R\} \quad (6.7)$$

For further details on this controversial issue, see for instance Refs [21], [28].

So we believe that this remark entitles us to re-consider the question of additivity of fuzzy uncertainty, and to this end, we shall use the framework of interval analysis.

6.2 Additivity of U_c -Uncertainty

(i) Consider two fuzzy intervals $f : X \rightarrow [0, 1]$, $z \rightarrow L(z)$ and $g : Y \rightarrow [0, 1]$, $z' \rightarrow L'(z')$. We define the pair (f, g) as

$$(f, g) : X \times Y \rightarrow [0, 1] \times [0, 1], \quad (z, z') \rightarrow (L(z), L'(z')) \tag{6.8}$$

then the following relation holds

$$(P6) \quad U_c((f, g)) = U_c(f) + U_c(g) \tag{6.9}$$

Indeed, one then has

$$U_c((f, g)) = \frac{\int_0^1 \int_0^1 L^c(z)[L'(z')]^c \ln L(z)L'(z') dz dz'}{\int_0^1 \int_0^1 L^c(z)[L'(z')]^c dz dz'} \tag{6.10}$$

therefore the result. The same rationale applies to discrete fuzzy sets.

As we mentioned it in subsection 5.2, everything looks like as if we randomized the discrete set (z_1, z_2, \dots, z_n) with the probability distribution $(L_1/L, L_2/L, \dots, L_n/L)$, and the definition (6.8) is quite consistent with this point of view.

(ii) The following relation holds

$$(P7) \quad U_0(f \otimes g) = U_0(f) + U_0(g) \tag{6.11}$$

but despite this feature, we shall advocate for $U_1(f)$ rather than $U_0(f)$, as we have already pointed out, mainly for the sake of consistency with the concept of entropy as it has been derived by Shannon.

(iii) We then have the relation

$$(P8) \quad U_1(f \otimes g) = \frac{\int_0^1 L'(z)L(z) \ln L(z) dz}{\int_0^1 L'(z)L(z) dz} + \frac{\int_0^1 L'(z)L(z) \ln L'(z) dz}{\int_0^1 L'(z)L(z) dz} \tag{6.12}$$

$$= \alpha(f, g)U_1(f) + \beta(f, g)U_1(g) \tag{6.13}$$

where $\alpha(f, g)$ and $\beta(f, g)$ are two positive coefficients which depend upon f and g .

6.3 On the Sub-Additivity of U_c -Uncertainty

Let $f(x, y)$, $f_Y(x)$ and $f_X(y)$ be defined as in equations (6.4) and (6.5); then one may have $U_1(f) \leq U_1(f_X \otimes f_Y)$ or on the contrary $U_1(f) \geq U_1(f_X \otimes f_Y)$. This surprising looking result is quite acceptable as far as we are dealing with subjectivity. Indeed, possibility is nothing but an approach to introducing subjectivity in probability, and it is well known that subjectivity can destroy or on the contrary create information [19].

7. Further Measures of Fuzzy Uncertainty

7.1 Renyi Uncertainty of Order s of Fuzzy Sets

(i) We first bear in mind that the Renyi entropy of order s of a continuous function $f : R \rightarrow R$, $x \rightarrow f(x)$ on the domain Ω is defined by the expression (the subscript R holds for Renyi)

$$H_{R,s}(f(\cdot); \Omega) := -\frac{1}{1-s} \ln \frac{\int_{\Omega} |f'(x)|^s dx}{\int_{\Omega} |f'(x)| dx}, \quad s > 1 \quad (7.1)$$

and one has

$$\lim H_{R,s}(f(\cdot); \Omega) = H_1(f(\cdot); \Omega) \text{ as } s \downarrow 1$$

where $H_1(f(\cdot); \Omega)$ is expressed by (3.1). For further details see the Ref. [18].

(ii) *Renyi entropy of fuzzy sets.* By applying the rationale of section 4 to the equation (5.13), one obtains the Renyi entropy of order s of the fuzzy interval $f(x)$ in the form

$$U_{R,s}(f) := -\frac{1}{1-s} \ln \frac{\int_0^1 L^s(z) dz}{\int_0^1 L(z) dz}, \quad s > 1 \quad (7.2)$$

and one can show easily that

$$\lim U_{R,s}(f) = U_S(f) \quad \text{as } s \downarrow 1 \quad (7.3)$$

For a discrete fuzzy set, we shall derive the corresponding Renyi uncertainty in the form

$$U_{R,s}(f) := -\frac{1}{1-s} \ln \left[\frac{\sum_{i=1}^n L^s(z)}{\sum_{i=1}^n L(z)} \right] \quad (7.4)$$

7.2 Structural Entropy of Order s of Fuzzy Sets

Havrda and Charvat [8] introduced the so-called structural entropy of order s of a discrete random variable X defined by the probability distribution $(p_1, p_2, p_3, \dots, p_n)$ in the form

$$H^s(X) := \left(\sum_{i=1}^n p_i^s - 1 \right) / (e^{1-s} - 1), \quad s > 1 \quad (7.5)$$

where e here is associated with the natural logarithm (one could select 2^{1-s} as well).

When $X \in R^n$ is continuous with the probability density $p(x)$, one has

$$H^s(X) := \left(\int_{R^n} p^s(x) dx - 1 \right) / (e^{1-s} - 1) \quad (7.6)$$

Later [20] we derived the corresponding structural entropy of order s of a deterministic mapping $f: R^n \rightarrow R^n$, $x \rightarrow f(x)$, on the domain Ω , in the form

$$H^s(f(\cdot); \Omega) := \frac{1}{e^{1-s} - 1} \left[\frac{\int_{\Omega} |f'(x)| dx}{\int_{\Omega} |f'(x)|^s dx} - 1 \right] \quad (7.7)$$

where $f'(x)$ denotes the Jacobian determinant of $f(x)$.

This can be extended to fuzzy intervals as follows.

Structural uncertainty of order s of fuzzy sets. Let $f(x)$, $X \rightarrow [0, 1]$ denote the membership function of a fuzzy set; then the substitution $|f'(x)| \leftarrow L(z)$ in (7.7) directly yields the structural uncertainty of this fuzzy set in the form

$$U^s(f) := \frac{1}{e^{1-s} - 1} \left[\frac{\int_0^1 L(z) dz}{\int_0^1 L^s(z) dz} - 1 \right], \quad s > 1 \quad (7.8)$$

8. Concluding Remarks and Outlook

8.1 Comments on the Meaningfulness of the Approach

In the framework of interval analysis, the logarithms of the z -cuts of a fuzzy set are quite meaningful measures of uncertainty on a componentwise standpoint, but at first glance, it is not clear how these local uncertainties are aggregated to yield the uncertainty of the whole fuzzy set itself. It is the main contribution of the information theory of deterministic functions, to provide a general possible modelling for this aggregation process.

As a matter of fact, we have obtained a family of weighting distributions indexed by a parameter c , and by adapting results previously derived for deterministic mappings, one can easily show that c pictures how the density $\ln L(z)$ of uncertainty is scrutinized by the observer. $c = 0$ defines a uniform observation scanning while $c = 1$ corresponds to Shannon entropy of random variables. Klir [9] focused on the case $c = 0$, that is to say on $U_0(f)$, but on doing so, is he not too much optimistic? Indeed, one has the inequality $U_0(f) < U_1(f)$, and $U_c(f)$ is the amount of uncertainty involved in the fuzzy set f to the observer!! In other words, he systematically assumes that the fuzziness of the observer is not so bad!

In short, if we had to select between $U_0(f)$ and $U_1(f)$, and if we assume that the maximum entropy principle [13], [14] holds, that we should choose $U_1(f)$.

8.2 Suggestions for Further Research

Technical communication in the presence of fuzziness. In 1975 [16], we introduced the concept of complete or total entropy of discrete random variables, in order to obtain a modelling of information loss. Later [17] we utilized the same concept to derive a

unified approach to discrete and continuous Shannon entropies. Here, in the presence of fuzziness, this entropy reads

$$H_e(X) := - \sum_{i=1}^n p_i \ln [p_i / \exp\{U_1(f_i)\}] \quad (8.1)$$

where (p_1, p_2, \dots, p_n) is the probability distribution defining X , and $U_1(f_i)$ is the fuzzy entropy which characterizes the i th-state x_i of X .

Given a second random variable Y which takes on the values (y_1, y_2, \dots, y_m) with the probability distribution (q_1, q_2, \dots, q_m) , one has the conditional total entropy

$$H_e(X/Y) := \sum_{j=1}^m q_j H_e(X/y_j) \quad (8.2)$$

with

$$H_e(X/y_j) := - \sum_{i=1}^n p_{i/j} \ln p_{i/j} + \sum_{i=1}^n p_{i/j} U_{1,j}(f_i) \quad (8.3)$$

$$= H(X/y_j) + \sum_{i=1}^n p_{i/j} U_{1,j}(f_i) \quad (8.4)$$

One can define an effective mutual information in the form

$$I_e(X/Y) = H(X) - H_e(X/Y) \quad (8.5)$$

and then expand a meaningful theory of communication in the presence of fuzziness.

Distance between fuzzy sets. The problem of defining the distance of two fuzzy sets is of interest, for instance, in studies related to approximate reasoning, and a possible new approach, via information of deterministic functions, is as follows.

Let $f(x)$ and $g(x)$ be the membership functions of two fuzzy intervals in R , and define the local Kullback divergence

$$d(f, g; x) := g(x) \ln \frac{g(x)}{f(x)} + \frac{1-g(x)}{1-f(x)} \ln \frac{1-g(x)}{1-f(x)} \quad (8.6)$$

then one could consider a possible distance in the form

$$D_c(f, g) := \frac{\int_R e^{cd(f,g;x)} d(f, g; x) dx}{\int_R e^{cd(f,g;x)} dx}, \quad c \in R \quad (8.7)$$

Mutual information between fuzzy sets. According to Shannon, the mutual information $I(X, Y)$ between the continuous random variables $X = \{x, p(x)\}$ and $Y = \{y, q(y)\}$ is

$$I(X, Y) := \int_{R^2} r(x, y) \ln [r(x, y) / p(x)q(y)] dx dy \quad (8.8)$$

where $r(x, y)$ is the probability density of the pair (X, Y) . Clearly, $I(X, Y)$ is the expectation of the density $I(x, y)$ of mutual information, expressed as

$$I(x, y) := -\ln p(x) - \ln q(y) - (-\ln r(x, y)) \quad (8.9)$$

The amount of mutual information involved in (X, Y) for $X \in \Omega \subset R$ and $Y \in \Omega' \subset R$ is then

$$I(X, Y; \Omega, \Omega') = \int_{\Omega} \int_{\Omega'} r(x, y) \ln[r(x, y)/p(x)q(y)] dx dy \text{ over } \int_{\Omega} \int_{\Omega'} r(x, y) dx dy \quad (8.10)$$

These remarks could be extended to fuzzy sets as follows. Consider the pair $f(x, y)$, $f_Y(x)$ and $f_X(y)$ as defined by equations (6.4) and (6.5); then according to (6.15), the density of mutual information at the z -cut level would be

$$\ln [L_{f_X}(z)L_{f_Y}(z)/L_f(z)]$$

and the mutual information $I_c(f, g)$ would be

$$I_c(f, g) := \frac{\int_R L_f^c(z) \ln [L_{f_X}(z)L_{f_Y}(z)/L_f(z)] dz}{\int_R L_f^c(z) dz} \quad (8.11)$$

References

- [1] Aczel, J. and Daroczy, Z., "On measures of Information and their Characterizations", *Academic Press*, New York, 1975.
- [2] Arbib, M. A., Book Review, "Human Systems Management", Vol 4, pp 310-324, 1984.
- [3] Belis, M. and Guiasu, S., "A qualitative-quantitative measure of information in cybernetic system", *IEEE Trans. Inform. Theory*, IT-14, pp 593-594, 1968.
- [4] Bouchon, B., "Entropic models", *Cybernetics and Systems*, Vol 18, pp 1-13, 1987.
- [5] De Luca. A. and Termini, S., "A definition of a nonprobabilistic entropy in the setting of fuzzy events", *Information and Control*, Vol 20, pp 301-312, 1972.
- [6] De Luca. and Termini, S., "Entropy and energy measures of a fuzzy set", in *Advances in Fuzzy Set Theory and Applications*, M. M. Gupta et als (Eds), pp 321-338, North-Holland, 1979.
- [7] French, S., "Fuzzy decision analysis: some criticisms", *TIMS Studies in Management Sciences*, Vol 20, pp 29-44, 1984.
- [8] Havrda, J. and Charvat, F., "Quantification method of classification Processes: concept of structural α -entropy", *Kybernetika*, Vol 3, pp 30-35, 1967.
- [9] Higashi, M. and Klir, G., "Measures of uncertainty and information based on possibility distributions", *Intern. J. General Systems*, Vol 9, pp 43-58, 1982.
- [10] Hirota, H., "Ambiguity based on the concept of subjective entropy", in *Fuzzy Information and Decision Processes*, M.M. Gupta et als (Eds) pp 29-40, North-Holland, 1982.
- [11] Hisdal, E., "Infinite valued logic based on two-valued logic and probability." Part 1.1. Difficulties with the present-day fuzzy-set theory and their resolution in the TEE model, *Intern. J. Man-Machine Studies*, Vol 25, pp 89-111, 1986.
- [12] Hisdal, E., "Infinite valued logic based on two-valued logic and probability." Part 1.2: Different sources of fuzziness, *Intern. J. Man-Machine Studies*, Vol 25, pp 113-138, 1986.

- [13] Jaynes, E. T., "Information theory and statistical mechanics 1", *Physical Review*, Vol 106, pp 620-630, 1957.
- [14] Jaynes, E. T., "Information theory and statistical mechanics 2", *Physical Reviews*, Vol 108, pp 171-190, 1957.
- [15] Jumarie, G., "Further advances on the general thermodynamics of open systems via information theory." Effective entropy, negative information, *Intern. J. Systems Sc.*, Vol 6, pp 249-268, 1975.
- [16] Jumarie, G., *Subjectivity, Information, Systems. Introduction to a Theory of Relativistic Cybernetics*, Gordon and Breach, New York, London, 1986.
- [17] Jumarie, G., "A Minkowskian theory of observation." Application to uncertainty and fuzziness, *Fuzzy Sets and Systems*, Vol 24, pp 231-255, 1987.
- [18] Jumarie, G., "Further results on the information theory of deterministic functions and its applications to pattern recognition", *Annales des Télécommunications*, Vol 45, pp 66-88, 1990.
- [19] Jumarie, G., *Relative Information. Theories and Applications*. Springer Verlag Berlin, New York, 1990.
- [20] Jumarie, G., "Structural entropies of order s and of order (r, s) for deterministic functions", *Cybernetics and Systems* (to appear) 1990.
- [21] Jumarie, G., "A theory of information for vagues concepts. Outline of application to approximate reasoning", *Kybernetes*, (to appear), 1990.
- [22] Manes, E.G., Book Review, *Bull. New Series Amer., Math. Soc.*, Vol 7, pp 603-612, 1982.
- [23] Moore, R. E., *Interval Analysis*, Prentice Hall Inc. New Jersey, 1966.
- [24] Muir, A., "Fuzzy sets and probability", *Kybernetes*, Vol 10, pp 197-200, 1981.
- [25] Muir, A., "Probability versus possibility: A reconciliation?" in *Proceedings 2nd Napoli Meeting on "The Mathematics of Fuzzy Systems"*, A.G.S. Ventre (Ed) pp 89-96, 1985.
- [26] Muir, A., "Relations between probability and fuzzy theory", in *The Mathematics of Fuzzy Systems*, A.G.S. Ventre (Ed) pp 249-263, Verlag TUV Rheinland, 1986.
- [27] Ramer, A., "Concepts of fuzzy information measures on continuous domains", *Intern. J. General Systems*, Vol 17, Nos 3-4, pp 1-11, 1990.
- [28] Schumucker, K.J., *Fuzzy Sets, Natural Language, Computations and Risk Analysis*, Computer Science Press, Rockville, Maryland, 1984.
- [29] Toth, H., "From fuzzy-set theory to fuzzy set-theory: some critical remarks on existing concepts", *Fuzzy Sets and Systems*, Vol 23, pp 219-237, 1987.
- [30] Trillas, E., Alsina, E. and Valverde, L., Do we need max, min and $I - J$ in fuzzy set theory? in *Applied systems and Cybernetics*, G.E. Lasker (Ed) Vol VI, pp 2808-2813, Pergamon Press, 1981.
- [31] Yager, R. R., "A measurement-informational discussion of fuzzy union and intersection", *Intern. J. Man-Machine Studies*, Vol 11, pp 189-200, 1979.
- [32] Zadeh, L.A., "Probability measures of fuzzy events", *J. Mathematical Analysis and Applications*, Vol 23, pp 421-427, 1968.
- [33] Zeleny, M., "On the (ir)relevance of fuzzy set theory", *Human Systems Management*, Vol 4, pp 301-306, 1984.

Dept. of Mathematics and Computer Sc. Université du Québec à Montréal, P.O. Box 8888, St A, Montréal, QUE, H3C 3P8, Canada.