Measuring Fuzzy Uncertainty

Nikhil R. Pal and James C. Bezdek, *Fellow, IEEE*

***Abstract*—First, this paper reviews several well known measures of fuzziness for discrete fuzzy sets. Then new *multiplicative* and *additive* classes are defined. We show that each class satisfies five well-known axioms for fuzziness measures, and demonstrate that several existing measures are relatives of these classes. The multiplicative class is based on non-negative, monotone increasing concave functions. The additive class requires only non-negative concave functions. Some relationships between several existing and the new measures are established, and some new properties are derived. The relative merits and drawbacks of different measures for applications are discussed. A weighted fuzzy entropy which is flexible enough to incorporate subjectivness in the measure of fuzziness is also introduced. Finally, we comment on the construction of measures that may assess *all* of the uncertainties associated with a physical system.**

***Index Terms*—Additive Measures, Entropy, Fuzzy sets, Measures of Fuzziness, Multiplicative Measures.**

I. Introduction

FUZZY SETS play a significant role in many deployed systems because of their capability to model nonstatistical imprecision. Consequently, characterization and quantification of fuzziness are important issues that affect the management of uncertainty in many system models and designs. This paper deals with quantitative measures of fuzziness.

Suppose a die is thrown and you are asked to guess the top face. Your uncertainty about the outcome is attributed to *randomness*. The best way to approach this question might be to describe the status of the die in terms of a probability distribution on the six faces. Uncertainty that arises due to chance is called *Probabilistic Uncertainty* (PU).

To make the situation more complex, suppose an artificial vision system analyzes a digital image of the top face. Based on the evidence gathered, the system might suggest that the top face is either a 5 or 6, but cannot be more specific. This kind of uncertainty arises from limitations (for example, sensor resolution) of the evidence gathering system. Uncertainty in this second situation reflects ambiguity in specifying the *exact* solution, and is called Non-specificity by Yager; we prefer to use the alternate term *Resolutional Uncertainty* (RU). If we are certain that the top face is either a 5 or 6, this case involves only nonspecficity. More generally, the vision system might also supply a certainty factor with its information. For example, the sytem might suggest that the top face is either a 5 or a 6 with belief of 0.8. In this case uncertainty due to chance is also present, because the top face can take any value, so the system contains both PU and RU.

Finally, suppose you are asked to interpret the top face of the die as, say, *high* (or *low*). Here a third type of uncertainty appears due to linguistic imprecision or vagueness; this is called *Fuzzy Uncertainty* (FU). Fuzzy uncertainty differs from PU and RU because it deals with situations where set boundaries are not sharply defined. Probabilistic and resolutional uncertainties are not due to ambiguity about set-boundaries, but rather, about the belongingness of elements or events to crisp sets. Uncertainty due to fuzziness is sometimes related to probabilistic uncertainty. For example, in the die experiment the occurrence of a 6 supports the fuzzy event HIGH more than a 3 does, but there is still an element of chance about the outcome of a throw, so the system contains PU and FU. Moreover, it is clear that RU can also appear in this third case, so the most complex systems may exhibit all three types. This article is devoted primarily to measures of FU, but we will also discuss several composite measures of PU and FU.

Section II reviews many existing measures of fuzziness. In Section III we propose two classes of measures for quantifying the fuzziness in a discrete fuzzy set. We refer to these as the *multiplicative* and *additive* classes. Several existing measures are shown to fit into these two frameworks. The new classes are extended to fuzzy sets with infinite support. Consistency of the additive and multiplicative classes with Yager's view of fuzziness is discussed. A relationship between several existing and the new measures is established that may help users select an appropriate measure for a specific application. Section IV introduces the concept of *weighted fuzzy entropy,* and describes a typical application of it. Several attempts have been made to combine fuzzy and probabilistic uncertainties [1]–[3]; Section V is devoted to a short review and discussion of some of these composite measures. Measures of nonfuzzy uncertainty (PU, RU) and various attempts at combining them are investigated elsewhere [4]–[17]. From the previous paragraphs it seems safe to assert that, generally, a complex system might contain all three (PU, RU, and FU) types of uncertainty. This paper concludes with a discussion about aggregating measures of all three types of uncertainty.

II. Measures of Fuzziness

Let X be any set. A fuzzy set (FS) A in X is characterized by a *membership function* μ_A: $X \rightarrow [0, 1]$. The value $\mu_A(x)$ represents the *grade of membership* of x in A. When μ_A is valued in $\{0, 1\}$, it is the characteristic function of a crisp (i.e., nonfuzzy) set. We let $\mathcal{P}(X)$ denote the set of all fuzzy subsets of X.

Manuscript received November 25, 1992; revised November 10, 1993.
N. R. Pal is with the Machine Intelligence Unit, Indian Statistical Institute, Calcutta 700035, India.
J. C. Bezdek is with the Computer Science Department, University of West Florida, Pensacola, Florida 32514.
IEEE Log Number 9216351.

1063–6706/94$04.00 © 1994 IEEE

Let χ be a discrete random variable that takes values in $X = \{x_1, x_2, \ldots, x_n\}$ with probabilities $P = \{p_1, p_2, \ldots, p_n\}$; we call (X, P) a *discrete probabilistic framework*. The set of all fuzzy subsets of $X = \{x_1, x_2, \ldots, x_n\}$ is denoted by $\mathcal{P}_n(X)$. Let log denote logarithm to any base >1. When $A \in \mathcal{P}_n(X)$ we let the n membership *values* of μ_A be denoted as $\{\mu_i\} = \{\mu_A(x_i), \quad 1 \leq i \leq n\}$.

Zadeh's *complement* A_Z^c of A is defined as $\mu_{A_Z^c}(x) = 1 - \mu_A(x)$ for each $x \in X$. Let A and B be two fuzzy subsets of X. The *union* and *intersection* of A and B were defined by Zadeh as follows:

$$\mu_{A\cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} \forall x \in X$$

and

$$\mu_{A\cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \forall x \in X.$$

Membership functions of fuzzy sets are never unique. Different individuals might define various μ_A's for the same fuzzy set. For example, the membership function of a fuzzy set TALL defined by an American for Americans would probably be different from that defined by an Oriental for Orientals. Jumarie [34] gives a detailed discussion of this issue, and remarks that the definition of fuzzy sets refers neither to the observed sets nor the observation process. This led Jumarie to define the concept of *relativistic* fuzzy sets. The relativistic fuzzy set describes the fuzziness which is involved in a set A with respect (relative) to an observer. Fuzziness in the relativistic framework is beyond the scope of this paper; interested readers are referred to [18]. In this paper we concentrate on fuzziness as defined by Zadeh.

Measures of fuzziness estimate the *average* ambiguity in fuzzy sets in some well-defined sense. We begin by considering properties that seem plausible for such a measure. The fuzziness of a crisp set using any measure should be zero, as there is no ambiguity about whether an element belongs to the set or not. If a set is maximally ambiguous ($\mu_A(x) = 0.5 \ \forall \ x$), then its fuzziness should be maximum. When a membership value approaches either 0 or 1, the ambiguity about belongingess of the argument in the fuzzy set decreases. A fuzzy set A^* is called a *sharpened* version of A if the following conditions are satisfied:

$$\mu_{A^*}(x) \leq \mu_A(x) \quad \text{if } \mu_A(x) \leq 0.5;$$

and

$$\mu_{A^*}(x) \geq \mu_A(x) \quad \text{if } \mu_A(x) \geq 0.5.$$

For a sharpened version A^* of A the measure of fuzziness should decrease because sharpening reduces ambiguity. Another intuitively desirable property is that the fuzziness measure of a set and its complement be equal. For example, the ambiguity present in the sets TALL and NOT TALL (note that NOT TALL is not necessarily SHORT) should be the same.

More formally, a measure of fuzziness for a discrete fuzzy set is a mapping $H\colon \mathcal{P}_n(X) \mapsto \Re^+$ that quantifies the degree of fuzziness present in A. Ebanks [19] suggested that measures of fuzziness should satisfy, for $A, B \in \mathcal{P}_n(X)$:

Sharpness $P1$: $H(A) = 0 \Leftrightarrow \mu_A(x) = 0$ or $1 \ \forall \ x \in X$;

Maximality $P2$: $H(A)$ is maximum $\Leftrightarrow \mu_A(x) = 0.5 \ \forall \ x \in X$;

Resolution $P3$: $H(A) \geq H(A^*)$, where A^* is a sharpened version of A.

Symmetry $P4$: $H(A) = H(1 - A)$, where $\mu_{1-A}(x) = 1 - \mu_A(x) \quad \forall \ x \in X$

Valuation $P5$: $H(A \cup B) + H(A \cap B) = H(A) + H(B)$.

Ebanks also proposed a sixth requirement called generalized additivity, but stated that it was somewhat difficult to interpret. We think $P1$–$P5$ are the correct requirements for measures of fuzziness. Ebanks gave the following necessary and sufficient conditions on functions that satisfy requirements $P1$–$P5$ for discrete fuzzy sets:

Theorem E [19, p. 28] Let $H\colon \mathcal{P}_n(X) \mapsto \Re^+$. Then H satisfies P1–P5 if and only if H has the form $H_E(A) = \sum_{i=1}^n g(\mu_i)$ for some function $g\colon [0, 1] \mapsto \Re^+$ that satisfies:

$G1$: $g(0) = g(1) = 0; \quad g(t) > 0 \forall t \in (0, 1)$;

$G2$: $g(t) < g(0.5) \forall t \in [0, 1] - \{0.5\}$

$G3$: *g is nondecreasing on $[0, 0.5)$ and nonincreasing on $(0.5, 1]$;*

$G4$: $g(t) = g(1 - t) \ \forall \ t \in [0, 1]$.

Theorem E provides a way to test measures of fuzziness against $P1$–$P5$, and is basic to our subsequent discussion. We call H_E an *E-Function* in the sequel.

The first attempt to quantify the uncertainty of a fuzzy set in the context of a discrete probablilistic framework apprears to have been made by Zadeh in 1968 [20], who defined the [weighted] entropy of $A \in \mathcal{P}_n(X)$ with respect to (X, P) as

$$H_{\mathrm{ZE}}(A, P) = -\sum_{i=1}^{n} \mu_i p_i \log\, p_i. \tag{1}$$

$H_{\mathrm{ZE}}(A, P)$ is a measure associated with fuzzy set A and the probabilities P that incorporates both probabilistic and fuzzy uncertainties. Equation (1) does not satisfy properties $P1$–$P5$, nor was it meant to. $H_{\mathrm{ZE}}(A, P)$ of a fuzzy event with respect to P is less than Shannon's entropy, $H_S(P) = -\sum_{i=1}^n p_i \log\, p_i$, of P alone.

The idea of measuring FU without reference to probabilities began in 1972 with the work of Deluca and Termini [1], who defined the entropy of $A \in \mathcal{P}_n(X)$ using Shannon's functional form,

$$H_{\mathrm{DTE}}(A) = -K \sum_{i=1}^{n} \mu_i \log \mu_i + (1 - \mu_i) \log(1 - \mu_i), \tag{2}$$

where K is a normalizing constant. Although (2) is an analog of Fermi's entropy, its meaning is quite different. Equation (2) satisfies properties $P1$–$P5$. Jumarie [18] comments that the second term, $(1 - \mu_i)\log(1 - \mu_i)$, is "theoretically" unnecessary; however, dropping this term causes H_{DTE} to fail $P2$. The sum in (2) can be factored into terms involving A and, A_Z^c

and this form has the alternate interpretation as the entropy of a fuzzy 2-partition of any n points. Bezdek generalized this in 1973 by defining the *classification entropy* of a constrained fuzzy c-partition U of $X = \{x_1, x_2, \ldots, x_n\}$ as $H_{\rm BE}(U) = -\sum_{k=1}^{n}\sum_{i=1}^{c}\mu_{ik}\log\mu_{ik}/n$ where $U = [\mu_{ik}]$ is a cxn matrix whose columns sum to 1, and whose rows are the membership functions of the c fuzzy clusters in X [21]. Bezdek also proposed the use of the Euclidean norm (called the *partition coefficient* in [21]), $H_B(U) = 1 - (\sum_{k=1}^{n}\sum_{i=1}^{c}\mu_{ik}^2/n)$, as a measure of the fuzziness of a constrained fuzzy c-partition U of X. Both of these measures apply to a single fuzzy set $A \in \mathcal{P}_n(X)$ by taking $c = 1$ in these formulae.

In the same year (1973), Capocelli and DeLuca [22] and subsequently DeLuca and Termini [23] initiated the study of fuzziness measures through the postulation of axioms such as properties $P1$–$P5$, and widened the specific form (2) to the first *class* of functions that satisfied intuitively desirable properties.

In 1975 Kaufmann [24] introduced an index of fuzziness for $A \in \mathcal{P}_n(X)$,

$$H_{\rm Ka}(A) = (2/n^k)d(A, A^{\rm near}), \tag{3}$$

where, $k \in R^+, d$ is a metric on $\mathcal{P}_n(X) \times \mathcal{P}_n(X)$, and $A^{\rm near}$ is the crisp set nearest to A (i.e., $A^{\rm near}(x) = \begin{Bmatrix} 1 & \text{if } \mu_A(x) \geq 0.5 \\ 0 & \text{otherwise} \end{Bmatrix}$. The value of k depends on d. For the (Minkowski) q-norms $d(A, A^{\rm near})$ and $H_{\rm Ka}(A)$ take the following forms:

$$d(A, A^{\rm near}) = \left[\sum_{i=1}^{n}|\mu_i - \mu_{A^{\rm near},i}|^q\right]^{1/q} \quad \text{so} \tag{4a}$$

$$H_{\rm Ka}(q, A) = \frac{2}{n^{1/q}}\left[\sum_{i=1}^{n}|\mu_i - \mu_{A^{\rm near},i}|^q\right]^{1/q} \quad \text{where} \tag{4b}$$

$q \in [1, \infty)$. $H_{\rm Ka}$ is called the *linear* or *quadratic* index of fuzziness as $q = 1$ or 2, respectively.

Kaufmann also mirrored Shannon's probabilistic entropy by using its form to define the entropy of a discrete fuzzy set. Letting $\varphi(\mu_i) = \mu_i/\sum_{j=1}^{n}\mu_j, i = 1, 2, \ldots, n$, Kaufmann's entropy of $A \in \mathcal{P}_n(X)$ is:

$$H_{{\rm Ka}E}(A) = -\frac{1}{\log(n)}\sum_{i=1}^{n}\varphi(\mu_i)\log\varphi(\mu_i). \tag{5}$$

Equation (5) does not depend directly on the membership values $\{\mu_i\}$; rather it is a function of their values relative to normalization by ϕ. This can cause an unexpected effect. For example, for the fuzzy sets with membership values $A = \{0.2, 0.2, \ldots, : n \text{ times}\}$ and $B = \{0.7, 0.7, \ldots, : n \text{ times}\}$, $H_{{\rm Ka}E}(A) = H_{{\rm Ka}E}(B) = 0$. Moreover $H_{{\rm Ka}E}$ does not satisfy properties $P2$–$P4$.

Knopfmacher [25] and Loo [26] have shown that (2) and (4b) are special cases of a larger class of measures defined as follows:

$$H_L(A) = h\left(\sum_{i=1}^{n}c_x g_x(\mu_i)\right), \tag{6}$$

where

i) $c_x \in R^+$ is a constant;
ii) $g_x : [0, 1] \to R^+$ is monotone increasing on [0, 0.5] and monotone decreasing on [0.5, 1];
iii) $g_x(0) = g_x(1) = 0$;
iv) g_x has a unique maximum at 0.5; and
v) $h : R^+ \to R^+$ is monotone increasing.

Equation (6) yields (2) (up to a constant) by defining g_x as $g_x(\mu_A(x)) = -\{\mu_A(x)\log\mu_A(x) + (1 - \mu_A(x))\log(1 - \mu_A(x))\}$ for all x, and h the identity function on R^+.

In 1978 Trillas and Riera [27] proposed a general class of fuzziness measures called $\otimes - *$ entropies which were defined as

$$H_{{\rm TR},\otimes-*}(A) = \otimes_{i=1}^{n}\alpha_i * N(\mu_i), \tag{7}$$

where for all i $\alpha_i > 0$ and N: $[0, 1] \mapsto \Re^+$ is an N-function that satisfies three properties:

$N1$. $N(0) = N(1) = 0$; otherwise, $N(t) > 0$;
$N2$. N is nondecreasing on [0, 0.5] and nonincreasing on [0.5, 1]; and
$N3$. N attains a maximum at 0.5.

In (7) $*$ is an inner operator that combines pairs $(\alpha_i, N(\mu_i))$; and $\otimes$ is an outer operator that aggregates these combinations. This class of entropies includes the important inner operators $* =$ minimum and $* =$ product; and outer operators $\otimes =$ sum and $\otimes =$ maximum. In particular $H_{\rm TR,sum-prod}(A) = \sum_{i=1}^{n}\alpha_i N(\mu_i)$ has the form of an E-function in Theorem E with $g(\mu_i) = \alpha_i N(\mu_i)$. For this choice of operators, properties $N1$–$N3$ of Trillas and Riera are equivalent to properties $G1$, $G3$ and $G2$ Ebanks, respectively. Recalling $G4$, we see that $H_{\rm TR,sum-prod}$ satisfies $P1$–$P5$ if and only if the N-function in $N1$–$N3$ satisfies the additional property $N(t) = N(1 - t)$ for all t in [0, 1]. Emptoz [28] also investigated properties of the sum-prod entropy.

Yager [29]–[30] associated fuzziness with the lack of distinction between a proposition and its negation. Yager used Zadeh's definition of complement to define his fuzziness measure:

$$H_Y(q, A) = (d^q(Y, Y^c) - d^q(A, A_Z^c))/d^q(Y, Y^c), \tag{8}$$

where $A \in \mathcal{P}_n(X), d^q$ is the q-norm having form (4a), and Y is an *arbitrary crisp* subset of X with complement Y^c. Note that $d^q(Y, Y^c)$ is the maximum distance between any pair of sets in $\mathcal{P}(X) \times \mathcal{P}(X)$.

Higashi and Klir [31] extended Yager's concept to a very general class of fuzzy complements. Let A^c be a *general* fuzzy complement function. Higashi and Klir introduced a class of measures based on aggregation of differences between the membership values in A and A^c. Let d aggregate the absolute differences $|\mu_i - \mu_i^c|$ in membership grades between A and A^c. Higashi and Klir's measure of fuzziness is

$$H_{\rm HK}(c, d; A) = \{d(Y, Y^c) - d(A, A^c)\}/d(Y, Y^c), \tag{9}$$

where Y is an arbitrary crisp subset of X. $d(Y, Y^c)$ is again the maximum distance in $\mathcal{P}(X) \times \mathcal{P}(X)$, and does not depend on the choice of crisp Y. The difference between Yager's measure (8) and the measure of Higashi and Klir (9) is in the complement and aggregation functions; (8) is a special cases of (9).

Kosko [32] defined a fuzziness measure as the ratio of the distance between the fuzzy set A and A^{near} to the distance between A and its *farthest* nonfuzzy set $A^{\text{far}} = (A^{\text{near}})^c$:

$$H_{\text{Ko}E}(q, A) = d^q(A, A^{\text{near}})/d^q(A, A^{\text{far}}), \tag{10}$$

where d^q is specified in (4a). $H_{\text{Ko}E}$ satisfies properties $P1$ through $P4$, but may not satisfy $P5$.

Pal and Pal [33] introduced a definition for probabilistic entropy based on an exponential gain function and used it as a basis for defining the following entropy of $A \in \mathcal{P}_n(X)$:

$$H_{\text{PPE}}(A) = K \sum_{i=1}^{n} \mu_i e^{1-\mu_i} + (1 - \mu_i)e^{\mu_i}, \tag{11}$$

where $e = 2.718\ldots$ and K is a constant. Equation (11) can be normalized so that it satisfies properties $P1$ through $P5$.

For differentiable membership functions Jumarie [18] suggested the following definition. Let $\mu' = \frac{d\mu}{dx}$, then the entropy of the fuzzy set $A \in \mathcal{P}(X)$ is defined as:

$$H_{\text{JE}}(A) = -\int_{-\infty}^{+\infty} |\mu'(x)| \log|\mu'(x)| dx. \tag{12}$$

Jumarie considered $|\mu'(x)|$ as the counterpart of a probability density and $\log|\mu'(x)|$ as the density of "uncertainty due to fuzziness". Regarding (12) we make the following observation: any probability density function $f(x)$ satisfies $\int_{-\infty}^{+\infty} f(x)dx = 1$, but this constraint is not satisfied by the fuzzy density function. As a result of this the integral in (12) can take *any* value! One solution to this problem would be to divide (12) by $\int_{-\infty}^{+\infty} |\mu'(x)|dx$, but even with this modification it is difficult to interpret H_{JE} as a measure of fuzziness.

Bhandari and Pal [35] defined a measure of fuzziness as the ratio of fuzzy divergence between A and A^{near} to that between A and A^{far}. The fuzzy divergence is an information measure for discrimination between two fuzzy sets. The definition of Bhandari and Pal satisfies properties $P1$–$P4$. Bhandari and Pal also use Renyi's entropy to define a nonprobabilistic entropy (a measure of fuzziness) for fuzzy sets as

$$H_{\text{BPE}}(\alpha : A) = \left(\frac{K}{1-\alpha}\right) \sum_{i=1}^{n} \log(\mu_i^\alpha + (1 - \mu_i)^\alpha), \tag{13}$$

where $\alpha > 0$, $\alpha \neq 1$ and K is a constant. Equation (13) satisfies properties $P1$–$P4$.

To study the average behavior of a fuzzy set, and to evaluate the extent to which μ separates the elements of X (compared to the ideal discrete case), a concept of separation and of the separating power of a fuzzy set was introduced by Dujet [36]. An entropy measure has also been suggested based on the concept of separation. Pal and Pal [2] introduced the concept of $r^{\text{th}}(r \geq 1)$ order fuzzy entropy, which gives the average ambiguity associated with any subcollection of r supports. Table I summarizes the measures of fuzziness (FU only) we have discussed so far; it also includes the two new ones that will be introduced in the next section (the last two rows of the table).

III. Two New Classes of Fuzziness Measures

Section II reviewed various measures of fuzziness, some of which satisfy one or more of properties $P1$–$P5$. In this section we propose two new classes—the *multiplicative* and *additive* classes of fuzziness measures—that satisfy these five properties.

A. The Multiplicative Class

Let $f : [0, 1] \to R^+$ be a concave increasing function on [0, 1]; that is, $f'(t) > 0$ and $f''(t) < 0 \forall t \in [0, 1]$. Now define

$$\hat{g}(t) = f(t)f(1 - t); \quad \text{and with } \hat{g}, \text{ define} \tag{14a}$$

$$g(t) = \hat{g}(t) - \min_{0 \leq t \leq 1}\{\hat{g}(t)\}. \tag{14b}$$

Finally, for $A \in \mathcal{P}_n(X)$, define

$$H_*(A) = K \sum_{i=1}^{n} g(\mu_i), \quad K \in R^+. \tag{15}$$

Since H_* has the form of an E-function as shown in Theorem E, the mulitiplicative class satisfies $P1$–$P5$ if and only if g in (14b) satisfies $G1$–$G4$. Theorem 1 confirms this result.

Theorem 1 H_* *satisfies P1–P5.*

Proof: Without loss take $K = 1$. Since, $\hat{g}(t) = f(t)f(1-t)$, $\hat{g}'(t) = f'(t)f(1 - t) - f(t)f'(1 - t)$. Now,

$$f'(t) > 0 \Rightarrow f(t) < f(1 - t) \text{ for } t \in [0, 0.5), \text{and} \tag{16a}$$

$$f''(t) < 0 \Rightarrow f'(1 - t) < f'(t) \text{ for } t \in [0, 0.5). \tag{16b}$$

Combining (16a) and (16b),

$$\hat{g}'(t) > 0 \text{ for } t \in [0, 0.5). \tag{17}$$

Similarly,

$$\hat{g}'(t) < 0 \text{ for } t \in (0, 5.1], \quad \text{and} \tag{18}$$

$$\hat{g}'(t) = 0 \text{ at } t = 0.5. \tag{19}$$

Combining (17)–(19), we see that $\hat{g}$ increases monotonically over [0, 0.5), attains a unique maximum at $t = 0.5$, and decreases monotonically over (0.5, 1], so $G2$ and $G3$ hold for $\hat{g}$. Since g and $\hat{g}$ differ by a constant, $G2$ and $G3$ also hold for g.

Since $\hat{g}$ takes minima at $t = 0$ and $t = 1$, $\hat{g}(0) = \hat{g}(1) = \min_{0 \leq t \leq 1} \{\hat{g}(t)\}$, so $g(0) = g(1) = 0$, and by monotonicity, $g(t) > 0$ for all $t \neq 0$ or 1, which is property $G1$. Equation (14a) shows that $\hat{g}$ satisfies $G4$, when $g(t) = \hat{g}(t) - \min_{0 \leq t \leq 1} \{\hat{g}(t)\} = \hat{g}(1 - t) - \min_{0 \leq t \leq 1} \{\hat{g}(1 - t)\} = g(1 - t)$, so g is symmetric with respect to $t = 0.5$, which is $G4$ for g. □

Example 1 Let

$$f : [0, 1] \to R^+ \text{ be defined as } f(t) = te^{1-t}. \tag{20}$$

Then $f'(t) = (1 - t)e^{1-t} > 0$ for $t \in [0, 1)$ and $f''(t) = (t - 2)e^{1-t} < 0$ for $t \in [0, 1]$, so f satisfies the requirements for the multiplicative class. Now $\hat{g}(t) = f(t)f(1 - t) =$

TABLE I
MEASURES OF FUZZY UNCERTAINTY (FU)

Author(s)	Measure of Fuzziness	Properties
DeLuca & Termini, 1972 (Entropy)	$H_{DTE}(A) = -K\sum_{i=1}^{n} \mu_i Log\mu_i + (1-\mu_i)Log(1-\mu_i)$	P1-P5
Bezdek, 1973 (Partition Entropy)	$H_{BE}(\{A_i\}) = -\sum_{i=1}^{c}(\sum_{k=1}^{n} \mu_{ik} Log\mu_{ik}/n)$ $A_i \in \mathcal{P}_n(X), i = 1,2,\ldots,c, \sum_{i=1}^{c}\mu_{ik} = 1\forall k$	NA
Kaufmann, 1975 (Index of fuzziness)	$H_{Ka}(q,A) = 2/n^{1/q}[\sum_{i=1}^{n} \vert\mu_i - \mu_i^{near}\vert^q]^{1/q}, q \in [1,\infty)$	P1-P4
Kaufmann, 1975 (Entropy)	$H_{KaE}(A) = -1/Log(n)\sum_{i=1}^{n}\varphi(\mu_i)Log\varphi(\mu_i)$	P1
Loo, 1977 (Fuzzy Uncertainty)	$H_L(A) = h(\sum_{i=1}^{n} c_x g_x(\mu_i))$	P1-P3
Trillas and Riera, 1978 ($\otimes - *$ Entropy)	$H_{TR,\otimes-*}(A) = \otimes_{i=1}^{n} \alpha_i * N(\mu_i)$	P1-P3
Yager, 1979 (Fuzzy Uncertainty)	$H_Y(q,A) = (d^q(Y,Y^c) - d^q(A,A_Z^c))/d^q(Y,Y^c)$, Y a crisp set	P1-P4
Higashi & Klir, 1982 (Fuzzy Uncertainty)	$H_{HK}(c,d;A) = \{d(Y,Y^c) - d(A,A^c)\}/d(Y,Y^c)$	P1-P4
Ebanks, 1983 (Fuzzy Uncertainty)	$H_E(A) = \sum_{i=1}^{n} g(\mu_i)$	P1-P5
Kosko, 1986 (Entropy)	$H_{KoE}(q,A) = d^q(A,A^{near})/d^q(A,A^{far})$	P1-P4
Pal & Pal, 1989 (Entropy)	$H_{PPE}(A) = K\sum_{i=1}^{n} \mu_i e^{1-\mu_i} + (1-\mu_i)e^{\mu_i}$	P1-P5
Jumarie, 1990 (Entropy)	$H_{JE}(A) = -\int_{-\infty}^{+\infty} \vert\mu'(x)\vert Log\vert\mu'(x)\vert\, dx$	NA
Bhandari & Pal,1992 (Entropy)	$H_{BPE}(\alpha : A) = (K/(1-\alpha))\sum_{i=1}^{n} Log(\mu_i^\alpha + (1-\mu_i)^\alpha)$	P1-P4
Pal and Bezdek, 1992 (Multiplicative FU)	$H_*(A) = K\sum_{i=1}^{n} g(\mu_i), K \in R^+, \hat{g}(t) = f(t)f(1-t)$ $g(t) = \hat{g}(t) - \min_{0\leq t\leq 1}\{\hat{g}(t)\}$, f concave & increasing	P1-P5
Pal and Bezdek, 1992 (Additive FU)	$H_+(A) = K\sum_{i=1}^{n} g(\mu_i), K \in R^+, \hat{g}(t) = f(t) + f(1-t)$ $g(t) = \hat{g}(t) - \min_{0\leq t\leq 1}\hat{g}(t)$, f concave	P1-P5

$te^{1-t}(1-t)e^t = e\ t(1-t)$, so $\min_{0\leq t\leq 1}\{\hat{g}(t)\} = 0$ and $g = \hat{g}$. Therefore, our construction yields the measure

$$H_{*\mathrm{QE}}(A) = K\sum_{i=1}^{n}\mu_i(1-\mu_i). \tag{21}$$

The constant e has here been absorbed in the constant K. It is easy to show directly that properties $P1$–$P5$ are satisfied by (21), as guaranteed by Theorem 1. For a constrained fuzzy c-partition U of X, it is easy to check that $H_{*\mathrm{QE}}(U) = nH_B(U)$ if $K = 1$. We call $H_{*\mathrm{QE}}$ the *quadratic entropy* of the fuzzy set because of its similarity to Vajda's probabilistic quadratic entropy for a discrete probabilistic framework,

$$H_{\mathrm{VQE}}(P) = \sum_{i=1}^{n} p_i(1-p_i). \tag{22}$$

Quadratic entropy was probably first used in theoretical physics by Fermi. It has also been used for risk evaluation of the nearest neighbor classification rule [37]. The choice of f in (20) was motivated by the expression for exponential entropy given in [33].

Example 2 Let $f : [0,1] \to R^+$ be $f(t) = t^\alpha, 0 < \alpha < 1$ so that $\min_{0\leq t\leq 1}\{\hat{g}(t)\} = 0$ and hence, $g = \hat{g}$. Now $f(t) \geq 0$; and $f'(t) = \alpha t^{\alpha-1} \geq 0$ and $f''(t) = \alpha(\alpha-1)t^{\alpha-2} < 0$. Since f satisfies the required conditions for the multiplicative class, the function

$$H_{\alpha\mathrm{QE}}(\alpha, A) = K\sum_{i=1}^{n}\mu_i^\alpha(1-\mu_i)^\alpha \tag{23}$$

is a measure of fuzziness satisfying $P1$–$P5$ which we shall call the *α-Quadratic entropy* for each $\alpha \in (0,1)$. We shall prove below that $H_{\alpha\mathrm{QE}} \to H_{*\mathrm{QE}}$ as $\alpha \xrightarrow{-} 1$. We discuss the effect and choice of α on $H_{\alpha\mathrm{QE}}$ after normalizing it with the constant shown in (24a) so that it lies in the range [0, 1]. Set

$$\hat{H}_{\alpha\mathrm{QE}}(\alpha, A) = \frac{1}{n2^{-2\alpha}}\sum_{i=1}^{n}\mu_i^\alpha(1-\mu_i)^\alpha$$

$$= \frac{1}{n}\sum_{i=1}^{n} S_{\alpha\mathrm{QE}}(\alpha, \mu_i), \text{ where} \tag{24a}$$

$$S_{\alpha\mathrm{QE}}(\alpha, \mu_i) = \frac{\mu_i^{\alpha}(1-\mu_i)^{\alpha}}{2^{-2\alpha}}. \tag{24b}$$

Theorem 2 For $\hat{H}_{\alpha\mathrm{QE}}$ as shown in (24), *we have:*

(A) $\mathrm{Lim}_{\alpha \xrightarrow{+} 0} \hat{H}_{\alpha\mathrm{QE}}(\alpha, A) = 1$;

(B) $\mathrm{Lim}_{\alpha \xrightarrow{-} 1} \hat{H}_{\alpha\mathrm{QE}}(\alpha, A) = \hat{H}_{*\mathrm{QE}}(A)$, *where $\hat{H}_{*\mathrm{QE}}$ is normalized at* (33a); *and*

(C) $\alpha_1 \le \alpha_2 \Rightarrow \hat{H}_{\alpha\mathrm{QE}}(\alpha_1, A) \ge \hat{H}_{\alpha\mathrm{QE}}(\alpha_2, A)$.

Proof: Verification of the limits in (A) and (B) is straightforward. For example, limit (A) is secured by observing that $\forall i, \mu_i^{\alpha} \to 1$, $(1-\mu_i)^{\alpha} \to 1$ and so $\mu_i^{\alpha}(1-\mu_i)^{\alpha} \to 1$ as $\alpha \to 0^+$. Further $\frac{1}{n2^{-2\alpha}} \to \frac{1}{n}$ with $\alpha \to 0^+$, so summing $\mu_i^{\alpha}(1-\mu_i)^{\alpha}$ over $i = 1$ to n yields (A). Even if $\mu_i = 0 \; \forall \; i$ $\mathrm{Lim}_{\alpha \xrightarrow{+} 0} \hat{H}_{\alpha\mathrm{QE}}(\alpha, A) = 1$. Do not interpret this as $\hat{H}_{\alpha\mathrm{QE}} = 1$ even when $\mu_i = 0 \; \forall \; i$. The fuzziness measure is defined for $0 < \alpha < 1$. For part (C), we calculate:

$$\begin{aligned}
&\hat{H}_{\alpha\mathrm{QE}}(\alpha_1, A) - \hat{H}_{\alpha\mathrm{QE}}(\alpha_2, A)\\
&= \frac{2^{2\alpha_1}}{n}\sum_i \{\mu_i(1-\mu_i)\}^{\alpha_1} - \frac{2^{2\alpha_2}}{n}\sum_i \{\mu_i(1-\mu_i)\}^{\alpha_2}\\
&= \frac{1}{n}2^{2\alpha_1}\sum_i \{\mu_i(1-\mu_i)\}^{\alpha_1}[1 - 2^{2(\alpha_2-\alpha_1)}\{\mu_i(1-\mu_i)\}^{\alpha_2-\alpha_1}].
\end{aligned}$$

Now

$$\begin{aligned}
\{\mu_i(1-\mu_i)\}^{\alpha_2-\alpha_1} &\le (1/4)^{\alpha_2-\alpha_1} \because \mu_i(1-\mu_i) \le 1/4\\
&\Rightarrow 2^{2(\alpha_2-\alpha_1)}\{\mu_i(1-\mu_i)\}^{\alpha_2-\alpha_1}\\
&\le 2^{2(\alpha_2-\alpha_1)}2^{-2}(\alpha_2-\alpha_1) = 1\\
&\Rightarrow 1 - 2^{2(\alpha_2-\alpha_1)}\{\mu_i(1-\mu_i)\}^{\alpha_2-\alpha_1} \ge 0\\
&\Rightarrow \hat{H}_{\alpha\mathrm{QE}}(\alpha_1, A) - \hat{H}_{\alpha\mathrm{QE}}(\alpha_2, A) \ge 0.
\end{aligned}$$

□

Fig. 1 depicts the function $S_{\alpha\mathrm{QE}}$ at (24b) for different values of α. When α is close to zero, $S_{\alpha\mathrm{QE}}$, and hence $\hat{H}_{\alpha\mathrm{QE}}$, does not change much with μ_i, unless μ_i is close to 0 or 1. On the other hand, when α is close to 1, each term of $S_{\alpha\mathrm{QE}}(\mu_i)$ behaves in the opposite fashion, and $\hat{H}_{\alpha\mathrm{QE}}$ is close to the quadratic entropy. Therefore, a choice of α around 0.5 will balance these two opposing tendencies. If α is much smaller than 0.5 then $\hat{H}_{\alpha\mathrm{QE}}$ will be insensitive to changes in μ, while if α is much larger than 0.5, $\hat{H}_{\alpha\mathrm{QE}}$ will be too sensitive to changes in μ.

The Additive Class For the additive class f is less restricted than for the multiplicative case. We require only that $f : [0,1] \to R^+$ be concave, $f''(t) < 0$. For the additive class the functions g and H_+ are defined as follows:

$$\hat{g}(t) = f(t) + f(1-t); \quad \text{and} \tag{25a}$$

$$g(t) = \hat{g}(t) - \min_{0 \le t \le 1}\{\hat{g}(t)\}. \tag{25b}$$

Finally, for $A \in \mathcal{P}_n(X)$, define

$$H_+(A) = K\sum_{i=1}^{n} g(\mu_i), \quad K \in R^+. \tag{26}$$

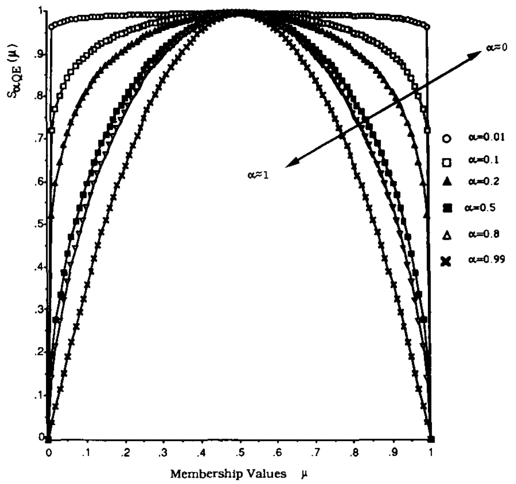


Fig. 1. Variation of $S_{\alpha\mathrm{QE}}$ with α.

Since H_+ in (26) also has the form shown in Theorem E, the additive class satisfies $P1$–$P5$ if g at (25b) satisfies $G1$–$G4$. Theorem 3 demonstrates this fact.

Theorem 3 H_+ satisfies P1–P5.

Proof: $\hat{g}'(t) = f'(t) - f'(1-t)$, and $f''(t) < 0 \Rightarrow f'(t) > f'(1-t)$ for $t \in [0, 0.5) \Rightarrow \hat{g}'(t) > 0$ for $t \in [0, 0.5)$. Similarly, $\hat{g}'(t) < 0$ for $t \in (0.5, 1]$, and $\hat{g}'(0.5) = 0$. Moreover, $\hat{g}(0) = \hat{g}(1)$. Thus $\hat{g}$ increases monotonically over [0, 0.5), attains its unique maximum at $t = 0.5$, and decreases monotonically over (0.5, 1]. Since g and $\hat{g}$ differ by a constant, g has the same behavior, and arguments similar to the multiplicative class can be used to establish $P1$–$P5$ for H_+. □

Example 3 Let $f : [0,1] \to R^+$ be $f(t) = te^{1-t}$, as in Example 1, (20). Now $\hat{g}(0) = \hat{g}(1) = 1$. Comparing H_+ to H_{PPE} at (11) for this f shows that $H_+(A) = H_{\mathrm{PPE}}(A) - n$, up to adjustment of K.

Example 4 Let $f : [0,1] \to R^+$ be defined as

$$f(t) = at - bt^2, 0 < b \le a. \tag{27}$$

Then $f(t) \ge 0$, and $f''(t) = -2b < 0$. Now $\hat{g}(0) = \hat{g}(1) = a - b$ so $g(t) = \hat{g}(t) + (b - a)$. Using (25) and (27) $g(t)$ has the form:

$$g(t) = 2bt(1-t). \tag{28}$$

Comparing H_+ to the quadratic fuzzy entropy at (21) shows that $H_+(A) = H_{*\mathrm{QE}}(A)$, subject to adjustment of the multiplicative constant. If $b = a$, then $\hat{g}(0) = \hat{g}(1) = 0$ and $g = \hat{g}$.

Example 5 Let $f : [0,1] \to R^+$ be defined as

$$f(t) = -t \log t. \tag{29}$$

Then $f(t) \ge 0$ and $f''(t) = -1/t < 0$ for $t \in (0, 1]$. Now $\hat{g}(0) = \hat{g}(1) = 0$, so $g = \hat{g}$. Comparing this H_+ with (2) shows that it becomes the logarithmic entropy of Deluca and Termini at (2), $H_+(A) = H_{\mathrm{DTE}}(A)$.

Extension to Membership Functions on Real Intervals The multiplicative and additive classes are easily extended to fuzzy sets on any real interval. Let A be a fuzzy subset of $X \subset R$. Then we define a measure of fuzziness (under the multiplicative class) for A as:

$$H_{*C}(A) = K \int_X g(\mu_A(x))dx, \tag{30}$$

where $K \in R^+$ is a constant and g is a continuous function whose form is defined in (14). H_{*C} reduces to H_* at (15) whenever X is discrete.

Theorem 4 H_{*C} *at (30) satisfies P1–P4*

Proof: We know that μ_A: $X \to [0, 1]$ and $\hat{g}(t)$ is a strictly concave function because $\hat{g}''(t) = f''(t)f'(1-t) - 2f'(t)f'(1-t) + f'(t)f''(1-t) < 0$. $\hat{g}(t)$ also symmetric about $t = 0.5$. In other words, $\hat{g}(t)$ attains the maximum value at $t = 0.5$ and a minimum value at $t = 0$ or 1. Therefore, $\hat{g}(\mu_A(x))$ attains its unique global maximum at $\mu_A(x) = 0.5$ and attains minimum values at $\mu_A(x) = 0$ or 1. Since $\hat{g}$ and g differ by a constant, the same statements are true for g. Hence, the integral in (30) will attain the global minimum value when $\mu_A(x) = 0$ or $1 \; \forall \; x \in X$. This proves $P1$. Similarly, $P2$ is also satisfied by H_{*C}, as it attains the unique global maximum when $\mu_A(x) = 0.5 \; \forall \; x \in X$. To prove $P3$, note that $g(\mu A^*(x)) \leq g(\mu_A(x)) \; \forall \; x \in X$, where A^* is a sharpened version of A. $P4$ is true by the definition of g. □

The additive class can also be extended as follows:

$$H_{+c}(A) = K \int_X g(\mu_A(x))dx, \tag{31}$$

where $K \in R^+$ is a constant and g is a continuous function whose form is defined in (25).

Theorem 5 H_{+c} *in* (31) *satisfies P1–P4.*

Proof: $\hat{g}''(t) = f''(t) + f''(1-t) < 0$ so $\hat{g}$ in (31) is strictly concave over [0, 1], and so g is as well. The rest of the proof is the same as Theorem 4. □

Remarks The $\hat{g}$-functions in our multiplicative and additive classes are generally not N-functions in the sense of Trillas and Riera [27] because they do not satisfy boundary condition $N1$ unless $\hat{g}(0) = \hat{g}(1) = 0$. Consequently, our additive class is somewhat different than $H_{\mathrm{TR,sum-prod}}$, but is closely related to it. Indeed, if f in (25a) is *strictly* concave and satisfies the conditions $f(0) = f(1) = 0$, then $g = \hat{g}$ in (25b). For these fs, H_+ becomes the entropy of DeLuca and Termini [23], which is in turn a special case of $H_{\mathrm{TR,sum-prod}}$. Theorems 4 and 5 have alternate derivations using results first given by Emptoz [28].

Consistency with Yager's View of Fuzziness Yager [29], [30] viewed fuzziness as a lack of distinction between a fuzzy set and its Zadeh-complement. The additive and multiplicative classes are consistent with this view, since the complement of any argument of g is $\mu_{A_z^c}(x) = 1 - \mu_A(x)$. Now $g(\mu_A(x))$ increases as $\mu_A(x)$ and $\mu_{A_z^c}(x)$ become closer. In other words, as the distinction between $\mu_A(x)$ and $\mu_{A_z^c}(x)$ decreases, $g(\mu_A(x))$ increases monotonically. Therefore, $g(\mu_A(x))$ can be viewed as a measure of lack of distinction between the values $\mu_A(x)$ and $\mu_{A_z^c}(x)$. Hence, H_+, H_{+c}, H_* and H_{*c} measure the overall lack of distinction between a fuzzy set and its Zadeh complement.

Comparative Analysis of Behavior First, we comment on the computational complexity of several of these measures. Equations (2), (4b) with $q > 1$, (5), (8)–(11), and (13) require either logarithmic or exponential evaluations, which are computationally more expensive than the multiplications and additions needed for quadratic entropy (21). Equations (2), (4b) with $q > 1$, (5), (8)–(11), and (13) also require additions and/or multiplications that are comparable to the total computational overhead of quadratic entropy.

Next we study the sensitivity of several measures under changes in the $\{\mu_i\}$. We consider only (2), (4b) with $q = 1$, (11), (13), (21), and (24). Each of these functions is normalized to have values in [0, 1] as follows:

$$\hat{H}_{\mathrm{Ka}}(q=1, A) = \frac{2}{n}\left(\sum_{i=1}^{n}\left|\mu_i - \mu_{\mathrm{near},i}\right|\right) = \frac{1}{n}\sum_{i=1}^{n} S_{\mathrm{Ka}}(\mu_i), \quad \text{where} \tag{32a}$$

$$S_{\mathrm{Ka}}(\mu_i) = 2 \operatorname{Min}\{\mu_i, 1-\mu_i\}. \tag{32b}$$

$$\hat{H}_{*\mathrm{QE}}(A) = \frac{4}{n}\left(\sum_{i=1}^{n}\mu_i(1-\mu_i)\right) = \frac{1}{n}\sum_{i=1}^{n} S_{*\mathrm{QE}}(\mu_i), \quad \text{where} \tag{33a}$$

$$S_{*\mathrm{QE}}(\mu_i) = 4\mu_i(1-\mu_i). \tag{33b}$$

$$\hat{H}_{\alpha\mathrm{QE}}(\alpha, A) = \frac{1}{n2^{-2\alpha}}\left(\sum_{i=1}^{n}\mu_i^{\alpha}(1-\mu_i)^{\alpha}\right) = \frac{1}{n}\sum_{i=1}^{n} S_{\alpha\mathrm{QE}}(\alpha, \mu_i), \quad \text{where} \tag{34a}$$

$$S_{\alpha\mathrm{QE}}(\alpha, \mu_i) = \frac{1}{2^{-2\alpha}}\{\mu_i^{\alpha}(1-\mu_i)^{\alpha}\}, \quad 0 < \alpha < 1 \tag{34b}$$

$$\hat{H}_{\mathrm{DTE}}(A) = -\frac{1}{n \log 2}\sum_{i=1}^{n}[\mu_i \log \mu_i + (1-\mu_i)\log(1-\mu_i)] = \frac{1}{n}\sum_{i=1}^{n} S_{\mathrm{DTE}}(\mu_i), \quad \text{where} \tag{35a}$$

$$S_{\mathrm{DTE}}(\mu_i) = -\frac{1}{\log 2}\{\mu_i \log \mu_i + (1-\mu_i)\log(1-\mu_i)\}. \tag{35b}$$

$$\hat{H}_{\mathrm{PPE}}(A) = \frac{1}{n(\sqrt{e}-1)}\sum_{i=1}^{n}\left(\mu_i e^{1-\mu_i} + (1-\mu_i)e^{\mu_i} - 1\right) = \frac{1}{n}\sum_{i=1}^{n} S_{\mathrm{PPE}}(\mu_i), \quad \text{where} \tag{36a}$$

$$S_{\mathrm{PPE}}(\mu_i) = \frac{\mu_i e^{1-\mu_i} + (1-\mu_i)e^{\mu_i} - 1}{(\sqrt{e}-1)}. \tag{36b}$$

$$\hat{H}_{\mathrm{BPE}}(\alpha : A) = \frac{1}{n(1-\alpha)\log 2} \times \left(\sum_{i=1}^{n}\log\left(\mu_i^{\alpha} + (1-\mu_i)^{\alpha}\right)\right) = \frac{1}{n}\sum_{i=1}^{n} S_{\mathrm{BPE}}(\alpha, \mu_i), \quad \text{where} \tag{37a}$$

$$S_{\mathrm{BPE}}(\alpha, \mu_i) = \frac{1}{(1-\alpha)\log 2} \log \{\mu_i^\alpha + (1-\mu_i^\alpha)\}. \qquad (37b)$$

Fig. 2 shows plots of one term (i.e., S) of the normalized measures in (32)–(37). For (34b) and (37b) $\alpha = 0.5$. We conclude from Fig. 2:

- ($S1$) there is not much difference in the overall behavior of one term of each measure except for S_{Ka}, the linear index of fuzziness;
- ($S2$) S_{Ka} is strongly and uniformly influenced by a small change in any membership value, so it is more sensitive to noise than the others; and
- ($S3$) the difference between one term of the normalized quadratic and exponential entropies, $S_{*\mathrm{QE}}$ and S_{PPE}, is neglible.

The normalized measures in (32)–(37) are composed of sums of terms having the general characteristics of the graphs in Fig. 2, so statements ($S1$)–($S3$) about each term are generally applicable to the measures themselves. The graphs in Fig. 2 suggest that each term of the six measures satisfies $S_{\mathrm{Ka}} \leq S_{*\mathrm{QE}} \leq S_{\mathrm{PPE}} \leq S_{\mathrm{DTE}} \leq S_{\alpha\mathrm{QE}}(\alpha = 0.5) \leq S_{\mathrm{BPE}}(\alpha = 0.5)$. In turn, this suggests that the same property holds for their sums (the $\hat{H}$s). We verify this in Theorem 6.

Theorem 6 $\hat{H}_{\mathrm{Ka}}(q = 1) \leq \hat{H}_{*\mathrm{QE}} \leq \hat{H}_{\mathrm{PPE}} \leq \hat{H}_{\mathrm{DTE}} \leq \hat{H}_{\alpha\mathrm{QE}}(\alpha = 0.5) \leq \hat{H}_{\mathrm{BPE}}(\alpha = 0.5)$.

Proof: Since each $\hat{H}$ is a separable function of the μ_i's, it suffices to show that $S_{\mathrm{Ka}}(t) \leq S_{*\mathrm{QE}}(t) S_{\mathrm{PPE}}(t) \leq S_{\mathrm{DTE}}(t) \leq S_{\alpha\mathrm{QE}}(\alpha = 0.5, t) \leq S_{\mathrm{BPE}}(\alpha = 0.5, t)$ for $t \in [0, 1]$.

To show $S_{\mathrm{Ka}}(t) \leq S_{*\mathrm{QE}}(t)$, consider $S_{*\mathrm{QE}}(t) - S_{\mathrm{Ka}}(t) = 4t(1-t) - 2\mathrm{Min}\{t, 1-t\} = \begin{cases} 4t(1-t) - 2t \text{ if } t \leq 0.5 \\ 4t(1-t) - 2(1-t) \text{ if } t \geq 0.5 \end{cases} = \begin{cases} 2t(1-2t) \text{ if } t \leq 0.5 \\ 2(1-t)(2t-1) \text{ if } t \geq 0.5 \end{cases} = \begin{Bmatrix} \geq 0 \text{ if } t \leq 0.5 \\ \geq 0 \text{ if } t \geq 0.5 \end{Bmatrix} \geq 0 \; \forall \; t \in [0, 1]$. Hence $S_{\mathrm{Ka}}(t) \leq S_{*\mathrm{QE}}(t)$. We prove that $S_{*\mathrm{QE}} \leq S_{\mathrm{PPE}} \leq S_{\mathrm{DTE}} \leq S_{\alpha\mathrm{QE}}(\alpha = 0.5) \leq S_{\mathrm{BPE}}(\alpha = 0.5)$ as follows. Each of $S_{*\mathrm{QE}}(t)$, $S_{\mathrm{PPE}}(t)$, $S_{\mathrm{DTE}}(t)$, $S_{\alpha\mathrm{QE}}(\alpha, t)$ and $S_{\mathrm{BPE}}(\alpha, t)$ is a *strictly* concave function for t in [0, 1] with unique maximum at $t = 1/2$ and minima at $t = 0, 1$. That is $S_{*\mathrm{QE}}(.5) = S_{\mathrm{PPE}}(.5) = S_{\mathrm{DTE}}(.5) = S_{\alpha\mathrm{QE}}(\alpha, .5) = S_{\mathrm{BPE}}(\alpha, .5) = 1$ and $S_{*\mathrm{QE}}(0 \text{ or } 1) = S_{\mathrm{PPE}}(0 \text{ or } 1) = S_{\mathrm{DTE}}(0 \text{ or } 1) = S_{\alpha\mathrm{QE}}(\alpha, 0 \text{ or } 1) = S_{\mathrm{BPE}}\,(\alpha, 0 \text{ or } 1) = 0$. Each of these functions is *symmetric* in t about $t = 0.5$. These facts enable us to assert that if $S_{*\mathrm{QE}}(t) \leq S_{\mathrm{PPE}}(t) \leq S_{\mathrm{DTE}}(t) \leq S_{\alpha\mathrm{QE}}\;(\alpha = 0.5, t) \leq S_{\mathrm{BPE}}\;(\alpha = 0.5, t)$ holds for any t in (0, 0.5), it must hold $\forall \; t \in (0, 0.5)$.

Without loss, at $t = 0.25$ we find $S_{*\mathrm{QE}}(0.25) = 0.75 \leq S_{\mathrm{PPE}}(0.25) = 0.7588 \leq S_{\mathrm{DTE}}(0.25) = 0.8113 \leq S_{\alpha\mathrm{QE}}\;(\alpha = 0.5, 0.25) = 0.866 \leq S_{\mathrm{BPE}}(\alpha = 0.5, 0.25) = 0.8999$. The result now follows by symmetry of the six functions. □

IV. Measures of Fuzziness in Image Processing

This section discusses fuzziness measures in the context of image processing. We noted in section 2 that the entropy of Zadeh [20] shown at (1) can be viewed as Shannon's entropy with each term weighted by the corresponding fuzzy membership. In the same year Belis and Guiasu [38] introduced a

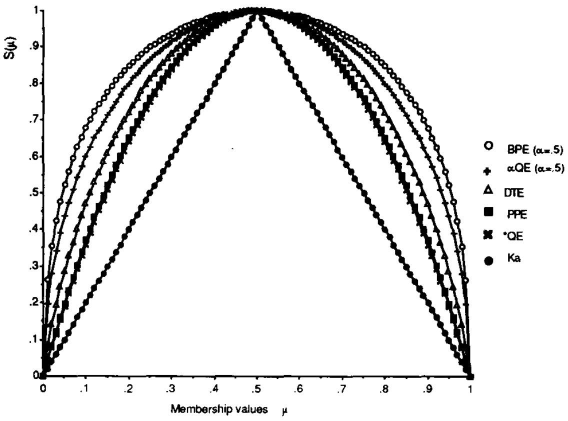


Fig. 2. Comparison of one term (S) of the six measures at (32)–(37).

slightly more general weighted Shannon entropy,

$$H_{\mathrm{BGE}}(P, W) = -\sum_{i=1}^{n} w_i p_i \log p_i, \qquad (38)$$

where $W = \{w_i \geq 0, i = 1, \ldots, n\}$ is a set of non-negative coefficients denoting the subjective importance of different outcomes of the discrete probabilistic framework (X, P). If all the weights are valued in [0, 1], (38) reduces to (1). Belis and Guiasu suggested that the occurrence of an event removes *two* types of uncertainty: the quantitative type related to its probability of occurrence, and the qualitative type relative to its utility (importance) for the fulfillment of some goal set by the experimenter. For example, depending on the goal, a low probability event might have high utility; while in a different context the same event might have low utility. If the weights satisfy the constraint $\sum_{i=1}^{n} w_i = 1$, their interpretation is more plausible. Belis and Guiasu suggested that W could be determined by a subjective goal set by the experimenter.

Weighted entropy allows subjective quantification of the information content associated with each event in (X, P). In this same spirit we introduce a similar concept for any discrete fuzzy set $A \in \mathcal{P}_n(X)$, by defining its *weighted fuzziness* as

$$H_{\mathrm{WFE}}(A, W) = \sum_{i=1}^{n} w_i g(\mu_i), \qquad (39)$$

where $W = \{w_i \geq 0 \; \forall \; i | \sum_i w_i = 1\}$, and g can have either the multiplicative (14b) or additive (25b) form. Sometimes the weights in (39) can be determined objectively. For example, when (X, P) is a discrete probability framework and $A \in \mathcal{P}_n(X)$, the expected fuzziness of outcomes in the system is found by taking $w_i = p_i$ in (39). In this case H_{WFE} does not measure the fuzziness associated with A as defined by $P1$–$P5$, nor it is equivalent to (1) or (38). Instead, H_{WFE} incorporates the subjectivity (discussed earlier) of fuzzy sets into our additive and multiplicative measures of fuzziness.

Expressions like (39) have been used for many applications [39]. For example, in [40] a digital image X with L gray levels was transformed into an array of fuzzy singletons, each with a membership value denoting the degree of brightness relative to some brightness level. To obtain a threshold for object-background segmentation the standard S-function was used. The S-function is defined by letting $\mathcal{L} = \{1, 2, \ldots, \mathcal{L}\}$ be the set of gray levels and $\mu_B:\mathcal{L} \to [0,1]$ be

$$\mu_B(l) = S(l; a, b, c) = \begin{cases} 0, & l \le a \\ 2\{(l-a)/(c-a)\}^2, & a \le l \le b \\ 1 - 2\{(l-c)/(c-a)\}^2, & b \le l \le c \\ 1, & l \ge c \end{cases},$$

with $b = (a+c)/2$. $\mu_B(l)$ can be interpreted as the membership of the gray level l in the fuzzy set WHITE (BRIGHT). Fig. 3 shows the plot of a typical S-function [41]. Note that $\mu_B(b) = 0.5$, as l increases above b membership in BRIGHT increases; conversely, as l decreases below b membership decreases.

The point b is called the cross-over point. If we define the brightness intensity of pixel (i,j) in X as $\mu_B(l_{ij})$ for $1 \le i \le M, 1 \le j \le N$, the resultant brightness image derived from X is a fuzzy subset $B \in \mathcal{P}_{\rm MN}(X)$ of X, i.e., a rectangular array of membership values $\{\mu_B(l_{ij}) = \mu_{ij}\}$. For a given window width $(c-a)$ and cross-over point b the entropy given in (2) was used in [40] for B,

$$H_{\rm DTE}(B) = -K \sum_{i=1}^{M} \sum_{j=1}^{N} \mu_{ij} \log \mu_{ij} + (1-\mu_{ij}) \log(1-\mu_{ij}). \quad (40)$$

Pal *et al.* [40] minimized $H_{\rm DTE}$ with respect to b to get a threshold for segmentation. In other words, (40) was evaluated for all possible integer cross-over points b, and b', the cross-over point that minimizes the entropy of the fuzzy image, was taken as the best threshold. Equation (40) is the *weighted* fuzzy entropy of BRIGHT. This can be seen by assuming that the L-level digital image is generated by an L-symbol source. Let p_i be the probability of gray level i in the image, so n_i/n is an estimate of p_i, where n is the total number of pixels in X and n_i is the number of times gray level i occurs the image. Rearranging (40),

$$\begin{aligned} H_{\rm DTE}(B) = -K' \sum_{i=1}^{L} [p_i(\mu_i \log \mu_i \\ + (1-\mu_i) \log(1-\mu_i))], \end{aligned} \quad (41)$$

where $K' = nK$. Thus, the expression used by Pal *et al.* in [40] is a weighted entropy.

This procedure for image segmentation can be implemented with *any* measure of fuzziness. In particular, $H_{\rm WFE}$ at (39) for the multiplicative and additive classes takes the forms:

$$\begin{aligned} H_{*\rm WFE}(W, B) = \sum_{i=1}^{L} w_i \{ f(\mu_i) f(1-\mu_i) \\ - \min_{0 \le t \le 1} \{ f(t) f(1-t) \} \}, \end{aligned} \quad (42a)$$

and

$$\begin{aligned} H_{*\rm WFE}(W, B) = \sum_{i=1}^{L} w_i \{ f(\mu_i) + f(1-\mu_i) \\ - \min_{0 \le t \le 1} \{ f(t) + f(1-t) \} \}, \end{aligned} \quad (42b)$$

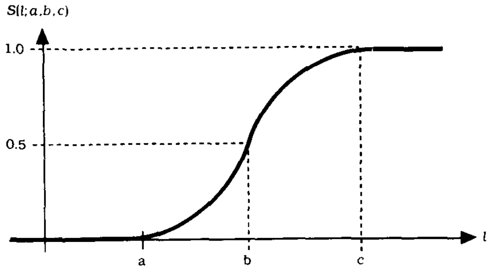


Fig. 3. Graph of the S function.

respectively. Either of equations (42) provides users a wide variety of choices because of freedom in the selection of the function f and the weight vector W. We emphasize that (40), (41), and (42) do not represent entropy (or any index of fuzziness) for the fuzzy set WHITE (BRIGHT), but the *weighted* fuzzy entropy of WHITE (BRIGHT).

V. Integration of Fuzzy and Probabilistic Uncertainties

There have been several attempts to combine probabilistic and fuzzy uncertainties when (X, P) is a discrete probability framework and $A \in \mathcal{P}_n(X)$ [1]–[3]. As mentioned in Section II, Zadeh [20] first defined the entropy of a fuzzy set with respect to a discrete probabilistic framework as the weighted Shannon entropy shown in (1). This entropy is a measure of uncertainty associated with a fuzzy event, and was the first composite measure of PU and FU. As seen in (1), $H_{\rm ZE}$ combines these two types of uncertainty with products involving both μ_i and p_i in each term of its sum.

Suppose we have a probabilistic framework (X, P), and there is some difficulty in interpreting x_i, the outcome of a trial, as 0 or 1. The average amount of ambiguity involved in the interpretation of such and outcome as suggested by Deluca and Termini [1] is:

$$H_{\rm DT}(A, P) = -\sum_{i=1}^{n} p_i(\mu_i \log \mu_i + (1-\mu_i) \log(1-\mu_i)). \quad (43)$$

$H_{\rm DT}$, like $H_{\rm ZE}$, combines PU with FU in each term in a nonseparable way. Deluca and Termini then defined the *total average uncertainty* in the system as the sum of Shannon's probabilistic entropy $H_S(P)$ and $H_{\rm DT}(A, P)$:

$$H_{\rm DT}^{\rm tot}(A, P) = H_S(P) + H_{\rm DT}(A, P). \quad (44)$$

$H_{\rm DT}^{\rm tot}$ is interpreted as the total average uncertainty incurred by making a prevision about the elements of X which appear as a result of the experiment, and in making a (0 or 1) decision about their values. Observe that both terms in (44) measure aspects of PU in different ways.

Xie and Bedrosian [3] defined the total entropy of a fuzzy set in a slightly different framework. Consider a crisp set $\hat{A}$ with only two kinds of elements, 0 and 1, having probabilities $P = \{p_0, p_1\}$, where $p_1 = 1 - p_0$. Suppose the sharpness of $\hat{A}$ is impaired, so that 0 takes some value in the interval [0, 0.5] and 1 takes some value in [0.5, 1]. Then $\hat{A}$ becomes a fuzzy set.

TABLE II
SOME COMPOSITE MEASURES OF PROBABILISTIC AND FUZZY UNCERTAINTY

Author(s)	Composite measures of fuzziness and probability
Zadeh, 1968 ([weighted] entropy)	$H_{ZE}(A,P) = \underbrace{-\sum_{i=1}^{n} \mu_i p_i Log\, p_i}_{PU+FU}$
Deluca and Termini, 1972 (average ambiguity)	$H_{DT}(A,P) = \underbrace{-\sum_{i=1}^{n} p_i(\mu_i Log\, \mu_i + (1-\mu_i) Log(1-\mu_i))}_{PU+FU}$
Deluca and Termini, 1972 (total average uncertainty)	$H_{DT}^{tot}(A,P) = \underbrace{H_S(P)}_{PU} + \underbrace{H_{DT}(A,P)}_{PU+FU}$
Xie and Bedrosian, 1984 (total uncertainty)	$H_{XB}^{tot}(A,P) = \underbrace{H_{DTE}(A)}_{FU} + \underbrace{H_S(P)}_{PU} \quad P = (p_0, p_1)$
Pal and Pal, 1992 (hybrid entropy)	$H_{PP}^{hyL}(A,P) = \underbrace{-p_0 Log\, E_0 - p_1 Log\, E_1}_{PU+FU}$,
	$E_0(A) = 1/n \underbrace{\sum_{i=1}^{n} (1-\mu_i) e^{\mu_i}}_{FU}$; $E_1(A) = 1/n \underbrace{\sum_{i=1}^{n} (\mu_i) e^{1-\mu_i}}_{FU}$
Pal and Pal, 1992 (hybrid exponential entropy)	$H_{PP}^{hyE}(A,P) = \underbrace{-p_0 e^{1-E_0} + p_1 e^{1-E_1}}_{PU+FU}$,
	$E_0(A) = 1/n \underbrace{\sum_{i=1}^{n} (1-\mu_i) e^{\mu_i}}_{FU}$; $E_1(A) = 1/n \underbrace{\sum_{i=1}^{n} (\mu_i) e^{1-\mu_i}}_{FU}$

Xie and Bedrosian [3] defined the total uncertainty associated with this system as the sum of Shannon's probabilistic entropy $H_S(P)$ and $H_{\mathrm{DTE}}(A)$, the fuzzy entropy of A at (2):

$$H_{\mathrm{XB}}^{\mathrm{tot}}(A,P) = H_S(P) + H_{\mathrm{DTE}}(A). \tag{45}$$

Equation (45) was perhaps the first composite measure of PU and FU that simply added together terms that measured each kind of uncertainty separately.

Pal and Pal [2] attempted to integrate probabilistic and fuzzy uncertainties by defining a *Hybrid entropy.* As in the previous case suppose p_0 and p_1 are the probabilities of 0 and 1 and μ_i denotes the membership of x_i in the fuzzy set $A \in \mathcal{P}_n(X)$ defined as "*symbols close to* 1". Pal and Pal then defined the hybrid entropy of A as

$$H_{\mathrm{PP}}^{\mathrm{hy}L}, (A,P) = -p_0 \log E_0 - p_1 \log E_1, \tag{46}$$

where E_0 and E_1 are defined as:

$$E_0 = \frac{1}{n}\sum_{i=1}^{n}(1-\mu_i)e^{\mu_i}; \quad \text{and} \tag{47a}$$

$$E_1 = \frac{1}{n}\sum_{i=1}^{n}(\mu_i)e^{1-\mu_i}. \tag{47b}$$

Expressions (47a) and (47b) are the two terms of the fuzzy entropy $H_{\mathrm{PPE}}(A)$ at (11). A detailed discussion on the justification for (46) can be found in [2]. Equation (46) was suggested by Shannon's probabilistic entropy. Pal and Pal also suggested another expression for hybrid entropy based on exponential entropy,

$$H_{\mathrm{PP}}^{\mathrm{hy}E}(A,P) = p_0 e^{(1-E_0)} + p_1 e^{(1-E_1)}, \tag{48}$$

with E_0 and E_1 as in (47).

The measures at (46) and (48), like (1) and (43), combine PU and FU in each of their terms.

We offer the following comments about integrating measures of fuzzy and probabilistic uncertainty. Since fuzziness is conceptually different from probabilistic uncertainty, algebraic summation of functions such as (45) that measure amounts of uncertainty due to each of these components is hard to justify. Referring back to the experiment of Deluca and Termini, in the absence of fuzziness, the system for (44) reduces to a two state system. Should the total uncertainty reduce to the probabilistic uncertainty of a two state system or an n-state system? Some justification can be given to favor *either* view. Moreover, (44) and (45) reduce to Shannon's entropy (for either an n-state system or a 2-state system) irrespective of the defuzzification process.

Measures such as (1), (43), (46) and (48) are even more difficult to interpret, because they combine PU and FU term by term. The hybrid entropies (46) and (48) reduce to the probabilistic (logarithmic or exponential) entropy of a 2-state system only with proper defuzzification. By proper defuzzification we mean the situation where the number of symbols (0 or 1) of each kind generated by the defuzzification process is the same as the number of symbols originally generated by the source. However, the hybrid entropies (46) and (48) and the total entropy (45) of Xie and Bedrosian cannot model the situation with an n-state system. Table II summarizes some of the composite measures discussed in this section.

VI. CONCLUSIONS

We have reviewed some well known measures of fuzziness, and introduced two general classes of measures for discrete fuzzy sets, the separable *additive* (H_+) and *multiplicative* classes (H_*). Fuzziness measures under each of these classes satisfy five well known axioms due to Ebanks. Defining a measure of fuzziness under either of these classes is easy, as the underlying functional form has few restrictions. The multiplicative class is based on non-negative, monotone increasing concave functions. The additive class is broader, requiring only non-negative concave functions. Several examples of each class were presented. We also demonstrated that several existing measures of fuzziness are related to these classes (up to additive and multiplicative constants). Theoretical properties of some existing and new measures were investigated. We also introduced a concept of weighted fuzzy entropy and discussed its power in incorporating subjectiveness in the measure of fuzziness.

Klir [42] argued that complex systems exhibit at least four types of uncertainty, and proposed somewhat different categories than ours. Specifically, he divided system uncertainty into two broad types; *fuzziness* due to vagueness; and *ambiguity*, which he further subdivided into nonspecificity, dissonance, and confusion. Our treatment (in [4],[5]) combines measures of dissonance and confusion under the heading of probabilistic uncertainty, because they are both generalizations of Shannon's entropy, and they are not independent.

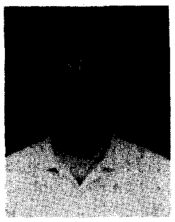
Attempts to integrate fuzzy and probabilistic uncertainties by adding components that measure each one have serious limitations because the relationship of these components is tenuous at best. Moreover, none of the functions considered here incorporates resolutional uncertainty (nonspecificity), nor even other kinds of uncertainty that may arise due to randomness (refer to the experiment of the die in the introduction). In the literature there are many definitions of entropy, many expressions for measures of fuzziness and many formulae for nonspecificity. Several expressions also exist for different aspects of uncertainty that may arise due to the presence of randomness in a system. Consideration of each facet of uncertainty separately, followed by injudicious combination of measures of them, leads to serious problems in interpretation and validation.

A much more meaningful approach would be to define a set of intuitively desirable axioms for the *total uncertainty in a system* due to randomness, fuzziness and nonspecificity, and then try to derive an expression that satisfies the axioms and quantifies it. This is hard. Although we have sets of axioms that seem desirable for each kind of uncertainty, it is difficult to conjecture a suitable set of axioms that leads to a satisfactory measure of the total uncertainty of a system. Hard, but important. This should be the topic of a future investigation.

REFERENCES

[1] A. Deluca and S. Termini, "A definition of nonprobabilistic entropy in the setting of fuzzy sets theory," *Inform. and Control*, vol. 20, pp. 301–312, 1972.
[2] N. R. Pal and S. K. Pal, "Higher order fuzzy entropy and hybrid entropy of a set," *Inform. Sci.*, vol. 61, no. 3, pp. 211–231, 1992.
[3] W. X. Xie and S. D. Bedrosian, "An information measure for fuzzy sets," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-14, pp. 151–156, 1984.
[4] N. R. Pal, J. C. Bezdek, and R. Hemasinha, "Uncertainty measures for evidential reasoning I: A review, "*Int. J. Approx. Reasoning*, vol. 7, pp. 165–183, 1992.
[5] ——, "Uncertainty measures for evidential reasoning II: A new measure," *Int. J. Approx. Reason.*, vol. 8, pp. 1–16, 1993.
[6] G. Shafer, *A mathematical theory of evidence*. Princeton, NJ: Princeton Univ. Press, 1976.
[7] R. R. Yager, "Entropy and specificity in a mathematical theory of evidence," *Int. J. Gen. Syst.*, vol. 9, pp. 249–260, 1983.
[8] U. Hohle, "Fuzzy plausibility measures," in *Proc. 3rd Int. Seminar on Fuzzy set theory*, E. P. Klement, Ed., Johannes Kepler Univ., Linz, 1981, pp. 7–30.
[9] U. Hohle, "Entropy with respect to plausibility measures," *Proc. 12th IEEE Int. Symp. Multiple Valued logic*, Paris, pp. 167–169, 1982.
[10] M. Higashi and G. J. Klir, "Measures of uncertainty and information based on possibility distributions," *Int. J. Gen. Syst.*, vol. 9, pp. 43–58, 1983.
[11] D. Dubois and H. Prade, "A note on measures of specificity for fuzzy sets," *Int. Jour. Gen. Syst.*, vol. 10, pp. 279–283, 1985.
[12] G. J. Klir and A Ramer, "Uncertainty in Dempster Shafer theory: A critical re-examination," *Int. J. Gen. Syst.*, vol. 18, pp. 155–166, 1990.
[13] D. Dubois and H. Prade, "A set theoretic view of belief functions: logical operations and approximations by fuzzy sets," *Int. J. Gen. Syst.*, vol. 12, pp. 193–226, 1986.
[14] R. R. Yager, "Measuring Tranquillity and anxiety in decision making: An application of fuzzy sets," *Int. J. Gen. Syst.*, vol. 8, pp. 139–146, 1982.
[15] M. T. Lamata and S. Moral, "Measures of entropy in the theory of evidence," *Int. J. Gen. Syst.*, vol. 14, pp. 297–305, 1987.
[16] D. Dubois and H. Prade, "Properties of measures of information in evidence and possibility theories," *Fuzzy sets and systems*, vol. 24, pp. 161–182, 1987.
[17] P. Smets, "Information content of an evidence," *Int. J. Man Mach. Studies*, vol. 19, pp. 33–43, 1983.
[18] G. M. Jumarie, *Relative information—theories and applications*. New York:Springer-Verlag, 1990.
[19] B. R. Ebanks, "On measures of fuzziness and their representations," *J. Math Anal. and Appl.*, vol. 94, pp. 24–37, 1983.
[20] L. A. Zadeh, "Probability measures of fuzzy events," *J. Math. Analysis and Applicat.*, vol. 23, pp. 421–427, 1968.
[21] J. C. Bezdek, "Fuzzy mathematics for pattern classification," Ph.D. Thesis, Cornell University, Ithaca, NY, 1973.
[22] R. M. Capocelli and A. Deluca, "Fuzzy sets and decision theory," *Inform. and Control*, vol. 23, pp. 446–473, 1973.
[23] A. Deluca and S. Termini, "Entropy of L-fuzzy sets," *Inform. and Control* vol. 24, pp. 55–73, 1974.
[24] A. Kaufmann, *Introduction to the theory of fuzzy subsets*. New York: Academic, vol. 1, 1975.
[25] J. Knopfmacher, "On measures of fuzziness," *J. Math. Anal. Applicat.*, vol. 49, pp. 529–534, 1975.
[26] S. G. Loo, "Measures of fuzziness, Cybernetica," vol. 20, pp. 201–210, 1977.
[27] E. Trillas and T. Riera, "Entropies in finite fuzzy sets," *Inform. Sci.*, vol. 15, pp. 159–168, 1978.
[28] H. Emptoz, "Nonprobabilistic entropies and indetermination measures in the setting of fuzzy sets theory," *Fuzzy Sets and Systems*, vol. 5, pp. 307–317, 1981.
[29] R. R. Yager, "On measures of fuzziness and negation, Part I: Membership in the unit interval," *Int. J. Gen. Syst.*, vol. 5, pp. 221–229, 1979.
[30] ——, "On measures of fuzziness and negation," *Part II: Lattices, Inform. and Contr.*, vol. 44, pp. 236–260, 1980.
[31] M. Higashi and G. J. Klir, "On measures of fuzziness and fuzzy complements," *Int. J. Gen. Syst.*, vol. 8, pp. 169–180, 1982.
[32] B. Kosko, "Fuzzy entropy and conditioning," *Inform. Sci.*, vol. 40, pp. 165–174, 1986.
[33] N. R. Pal and S. K. Pal, "Object-background segmentation using new definitions of entropy," *IEE Proc.*, 136, pt. E, pp. 284–295, 1989.
[34] G. M. Jumarie, *Subjectivity, information, systems—introduction to a theory of relativistic cybernetics*. New York: Gordon and Breach, 1986.

[35] D. Bhandari and N. R. Pal, "Some new information measures for fuzzy sets," *Inform. Sci.*, 1992 (to appear).
[36] C. Dujet, "Separation and measures of fuzziness," *Fuzzy Sets and Syst.*, vol. 28, pp. 245–262, 1988.
[37] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inform. Theory*, vol. IT–13, pp. 21–27, 1967.
[38] M. Belis and S. Guiasu, "A qualitative-quantitative measure of information in cybernetic system," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 593–594, 1968.
[39] S. K. Pal and D. Dutta Majumder, *Fuzzy Mathematical Approach to Pattern Recognition.* New York: Wiley, 1986.
[40] S. K. Pal, R. A. King, and A. A. Hasim, "Automatic grey level thresholding through index of fuzziness and entropy," *Pattern Recognit. Lett.*, vol. 1, pp. 141–146, 1983.
[41] L. A. Zadeh, K. S. Fu, K. Tanaka, and M. Shimura, Eds., *Fuzzy Sets and their Applications to Cognitive and Decision Processes.* London: Academic, 1975.
[42] G. J. Klir, "Where do we stand on measures of uncertainty, ambiguity, fuzziness and the like?" *Syst. Research*, vol. 2, pp. 131–140, 1985.



Nikhil R. Pal (M'91) received the B.Sc. degree with honors in physics and the Master of Busines Management degree from the University of Calcutta in 1979 and 1982, respectively. He received the M.Tech. (Comp. Sc.) and Ph.D. (Comp. Sc.) degrees from the Indian Statistical Institute in 1984 and 1991, respectively.

Currently, he is associated with the Machine Intelligence Unit of Indian Statistical Institute, Calcutta. From September 1991 to February 1993 he was with the Computer Science Department of the University of West Florida. He was guest faculty of the University of Calcutta also. His research interest includes image processing, fuzzy sets theory, measures of uncertainty, neural networks, and genetic algorithms. He is an associate editor of the *International Journal of Approximate Reasoning* and IEEE TRANSACTIONS ON FUZZY SYSTEMS.

Jim Bezdek (M'80–SM'90–F'92) received the Ph.D. degree from Cornell University, NY, in 1973.

His interests include pattern recognition, fishing, computational neural networks, skiing, image processing, blues music, medical computing, and motorcycles. He is the founding editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS.