International Journal of Fuzzy Mathematics and Systems.
ISSN 2248-9940 Volume 13, Number 1 (2023), pp. 1-9

http://www.ripublication.com

Some Fuzzy Noiseless Coding Theorems Connected with Non-Shannon Information Measures and their Bounds

Rohit Kumar Verma

Department of Mathematics
Bharti Vishwavidyalaya, Durg, Chhattisgarh, India

Abstract

In this discussion we attempt to prove the noiseless coding theorems on the basis of new generalization of fuzzy information measures. The results investigate not only for average code word length L_a^b but also for their bounds in fuzzy environments are the special cases.

Keywords: Holder inequality, Kraft inequality, Fuzzy noiseless coding, Fuzzy set, Decipherable code etc.

1. Introduction

Fuzziness and uncertainty are the basic nature of human thinking and many real world intents. The concept of fuzziness was first time introduced by Zadeh, Lotfi A. [17]. He developed a new theory to measure the uncertainty, which is also known as ambiguity, of a fuzzy set. This theory plays a significant role the ambience of information theory. If A be the subset of universe of discourse $i.e. X = \{x_1, \dots\dots, x_n\}$ then, A is defined as,

$$A = \{x_i/\mu_A(x_i): i = 1,2, \dots\dots\dots\dots, n\}.$$

Where $\mu_A(x_i)$ is a membership function and having the following properties:

1. If $\mu_A(x_i) = 0$, x_i does not belong to A and there is no ambiguity.
2. If $\mu_A(x_i) = 1$, x_i belong to A and there is no ambiguity.
3. If $\mu_A(x_i) = 0.5$, there is maximum ambiguity whether x_i belong to A or not.

By the definition, a fuzzy set is a set containing elements with varying membership degrees. This is different from classical (crisp) sets in which elements have full membership in that set.

If $x_1, x_2, \dots \dots, x_n$ are member of universe of discourse, then all $\mu_A(x_1), \mu_A(x_2), \dots, \mu_A(x_n)$ lies between 0 and 1, but these are not probabilities because their sum is not unity. $\mu_A(x_i)$, gives the degree of belongingness of the element x_i to the set A. The function $\mu_A(x_i)$ associates with each $x_i \in R^n$ a grade of membership to the set A and is known as membership function or finite fuzzy information scheme $i.e.$

$$F.S = \begin{bmatrix} x_1 \; x_2 \; \dots \; x_n \\ \mu_A(x_1) \; \mu_A(x_2) \dots \mu_A(x_n) \end{bmatrix},$$

$$0 \le \mu_A(x_i) \le 1 \text{ for all } x_i \in X. \tag{1.1}$$

Later, many other researchers increase this idea and made more efforts in this particular area. For instance, Kaufmann [6] proposed fuzzy entropy of a fuzzy set by a metric distance between its membership function and the membership function of its nearest crisp set.

The importance of fuzzy sets comes from the fact that it can deal with imprecise and mexact information, many fuzzy measures have been discussed and derived by Kapur [8], Lowen [9], Nguyen and Walker [11] etc. The basic noiseless coding theorems are developed by Aczel [1], Kapur [8], Renyi [12] and Van Der Lubbe [16]. They also developed the lower bound for the mean code-word length of a uniquely decipherable code in terms of Shannon's [13] entropy. Kapur [8] has established relationships between probability, entropy and coding.

Each and every finite scheme describes a state of ambiguity, corresponding to Shannon's [13] probabilistic entropy.

In 1972, De-Luca and Termini [4] proposed the measure of fuzzy entropy

$$H(A) = -\frac{1}{n}\sum_{i=1}^{n}\left[\mu_A(x_i)\ln\mu_A(x_i) + \left(1-\mu_A(x_i)\right)\ln\left(1-\mu_A(x_i)\right)\right]. \tag{1.2}$$

This measure serves as a very suitable measure of fuzzy entropy for the finite fuzzy information scheme. Similarly Kapur [8] suggests the fuzzy information measure corresponding to Havrda-Charvats [5] was as follows

$$H^a(A) = \frac{1}{1-a}\sum_{i=1}^{n}\left(\left(\mu_A^a(x_i) + \left(1-\mu_A(x_i)\right)^a\right) - 1\right),$$

$$a \neq 1, a > 0. \tag{1.3}$$

In section 2 we derive the average code word length connected with fuzzy entropy corresponding to Verma's [15] entropy, which is also connected with Havrda-Charvat [5], Tsallis [14], Shannon [13] and Mitter-Mathur's [10] entropy.

2. Our Results

Coding Theorems in Fuzzy Environment

In this part, we define fuzzy measures of information [15] as

$$V_a^b\big(\mu_A(x_i)\big) = \frac{1}{a-1}\left(1 - \frac{\sum_{i=1}^{n}\left(\mu_A^{ab}(x_i) + \big(1-\mu_A(x_i)\big)^{ab}\right)}{\sum_{i=1}^{n}\left(\mu_A^{b}(x_i) + \big(1-\mu_A(x_i)\big)^{b}\right)}\right) \tag{2.1}$$

Now, when $b \to 1$, (2.1) reduces into Havrda-Charvat [5] and Tsallis's [14] fuzzy measures of information $i.e.$

$$V_a\big(\mu_A(x_i)\big) = \frac{1}{1-a}\sum_{i=1}^{n}\left(\left(\mu_A^{a}(x_i) + \big(1-\mu_A(x_i)\big)^{a}\right) - 1\right) \tag{2.2}$$

and when $b = 1$ and $a \to 1$ then (2.1) reduces into Shannon's [13] fuzzy measures of information $i.e.$

$$V\big(\mu_A(x_i)\big) = -\sum_{i=1}^{n}\Big(\mu_A(x_i)\ln\mu_A(x_i) - \big(1-\mu_A(x_i)\big)\ln\big(1-\mu_A(x_i)\big)\Big) \tag{2.3}$$

Again, when $a \to 1$, (2.1) reduces into Mitter and Mathur's [10] fuzzy measures of information $i.e.$

$$V^b\big(\mu_A(x_i)\big) = -\frac{\sum_{i=1}^{n}\left(\mu_A^{b}(x_i)\ln\mu_A^{b}(x_i) + \big(1-\mu_A(x_i)\big)^{b}\ln\big(1-\mu_A(x_i)\big)^{b}\right)}{\sum_{i=1}^{n}\left(\mu_A^{b}(x_i) + \big(1-\mu_A(x_i)\big)^{b}\right)} \tag{2.4}$$

Theorem 2.1: For all uniquely decipherable code

$$V_a^b(A) \le L_a^b \tag{2.1.1}$$

under the condition $\sum_{i=1}^{n} D^{-l_i} \le 1$ for all $D > 1$, equality holds if and only if

$$l_i = -\ln_D\left(\frac{\mu_A^{ab}(x_i) + \big(1-\mu_A(x_i)\big)^{ab}}{\sum \mu_A^{ab}(x_i) + \big(1-\mu_A(x_i)\big)^{ab}}\right) \tag{2.1.2}$$

Proof: We know that for all $x_i, y_i > 0$, $i = 1, \dots\dots, n$ and $\frac{1}{\gamma} + \frac{1}{\delta} = 1$, $\gamma < 1(\neq 0)$, $\delta < 0$ or $\delta < 1(\neq 0)$, $\gamma < 0$, then the Holder's inequality

$$\left(\sum_{i=1}^{n} x_i^{\gamma}\right)^{\frac{1}{\gamma}} . \left(\sum_{i=1}^{n} y_i^{\delta}\right)^{\frac{1}{\delta}} \le \sum_{i=1}^{n} x_i . y_i \tag{2.1.3}$$

holds, and equality holds in (2.1.3) if and only if there exists a positive constant μ such that

$$x_i^{\gamma} = \mu y_i^{\delta} \tag{2.1.4}$$

Set

$$x_i = \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{-\frac{1}{t}} D^{-l_i}$$

and

$$y_i = \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{\frac{1}{t}}$$

Also,

$$\gamma = -t \Rightarrow 0 < \gamma < 1$$

and

$$\delta = \frac{t}{t+1} \Rightarrow \delta < 0.$$

Hence from equation (2.1.3), we achieve

$$\left[\left\{\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{-\frac{1}{t}} D^{-l_i}\right\}^{-t}\right]^{-\frac{1}{t}} \times \left[\left\{\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{\frac{1}{t}}\right\}^{\frac{t}{t+1}}\right]^{\frac{t+1}{t}}$$

$$\leq \sum_{i=1}^{n} \left[\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{-\frac{1}{t}} D^{-l_i} . \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{\frac{1}{t}}\right]$$

Now, applying Kraft's inequality, we get the result

$$\left[\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{\frac{1}{t+1}}\right]^{\frac{t+1}{t}} \leq \left[\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right] D^{l_i t}\right]^{\frac{1}{t}}$$

$i.e.$

$$\sum_{i=1}^{n} \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right]^{\frac{1}{t}} \leq \sum_{i=1}^{n} \left[\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right] D^{l_i t}\right]^{\frac{1}{t}}$$

$i.e.$

$$\sum_{i=1}^{n} \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right] \leq \sum_{i=1}^{n} \left[\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right] D^{l_i t}\right]$$

Multiplying both sides by t, we achieve

$$t \sum_{i=1}^{n} \left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right] \leq t \sum_{i=1}^{n} \left[\left[f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right] D^{l_i t}\right]$$

Subtracting n from both sides, we get

$$t\sum_{i=1}^{n}\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)-1\right]\leq t\sum_{i=1}^{n}\left[\left[f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right]D^{l_i t}-1\right] \quad (2.1.4)$$

Taking

$$a=\frac{t-1}{t}, t=\frac{1}{1-a}, a>0, a\neq 1$$

and

$$f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)=\frac{\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}{\sum\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}$$

$i.e.$

$$\frac{1}{1-a}\sum_{i=1}^{n}\left[\left\{\frac{\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}{\sum\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}\right\}-1\right]$$

$$\leq\frac{1}{1-a}\sum_{i=1}^{n}\left[\left\{\frac{\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}{\sum\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}\right\}D^{l_i t}-1\right]$$

$i.e.$

$$\frac{1}{a-1}\sum_{i=1}^{n}\left[1-\left\{\frac{\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}{\sum\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}\right\}\right]$$

$$\leq\frac{1}{a-1}\sum_{i=1}^{n}\left[1-\left\{\frac{\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}{\sum\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}}\right\}D^{l_i t}\right]$$

$i.e.$

$$V_a^b(A)\leq L_a^b$$

and the equality holds if and only if

$$D^{-l_i}=f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)$$

$i.e.$

$$l_i=-\ln_D\left(f\left(\mu_A(x_i),\mu_{A^c}(x_i)\right)\right)$$

where,

$$f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)=\frac{\mu_A^{ab}(x_i)+\big(1-\mu_A(x_i)\big)^{ab}}{\sum\mu_A^{ab}(x_i)+\big(1-\mu_A(x_i)\big)^{ab}}$$

Theorem 2.2: For every uniquely decipherable codes

$$V_a^b(A).D^{1-a}+\frac{1}{a-1}\{1-D^{1-a}\}>L_a^b$$

Proof: Suppose l_i be the code word length for all $i=1,\dots\dots,n$, which satisfies the condition

$$-\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)<l_i<-\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)+1 \qquad (2.2.1)$$

where

$$f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)=\frac{\mu_A^{ab}(x_i)+\big(1-\mu_A(x_i)\big)^{ab}}{\sum\mu_A^{ab}(x_i)+\big(1-\mu_A(x_i)\big)^{ab}}.$$

Suppose the unit interval be

$$\Delta_i=\Big[-\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big),-\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)+1\Big] \qquad (2.2.2)$$

Then, for any Δ_i, there exists exactly one positive number l_i such that

$$0<-\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)\le l_i<-\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)+1 \quad (2.2.3)$$

Since the sequence of length $\{l_i\}_{i=1}^n$ satisfies the condition $\sum_{i=1}^n D^{-l_i}\le 1$. So from (2.2.3) we achieve

$$l_i<-\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)+1$$

$i.e.$

$$-l_i>\ln_D\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)-1$$

$i.e.$

$$D^{-l_i}>f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big).D^{-1}$$

$i.e.$

$$D^{-l_i\left(\frac{a-1}{a}\right)}>\Big(f\big(\mu_A(x_i),\mu_{A^c}(x_i)\big)\Big)^{\frac{a-1}{a}}D^{\frac{1-a}{a}} \qquad (2.2.4)$$

Multiplying both sides of (2.2.4) by $\frac{\mu_A^b(x_i)+\left(1-\mu_A(x_i)\right)^b}{\left(\sum\mu_A^b(x_i)+\left(1-\mu_A(x_i)\right)^b\right)^{\frac{1}{a}}}$, summing over $i=$

$1, \ldots \ldots, n$ and simplification for $\frac{1}{a-1}$ as $a > 1$. Hence the desired result.

Theorem 2.3: For every code word length L_i, show that the inequality

$$(a-1)L_a^b \geq (a-1)V_a^b(A) > (a-1)V_a^b(A).D - (D-1)$$

Proof: Letting,

$$L_i = -\ln_D\left(f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right)\right)$$

Where,

$$f\left(\mu_A(x_i), \mu_{A^c}(x_i)\right) = \frac{\mu_A^{ab}(x_i) + \left(1-\mu_A(x_i)\right)^{ab}}{\sum \mu_A^{ab}(x_i) + \left(1-\mu_A(x_i)\right)^{ab}} \tag{2.3.1}$$

Ofcourse, L_i and $L_i + 1$ satisfy the equality in Holder's inequality. On the other hand L_i satisfies the condition $\sum_{i=1}^n D^{-L_i} \leq 1$. Also, l_i satisfies Kraft's inequality $i.e. \sum_{i=1}^n D^{-l_i} \leq 1$ because l_i lies between L_i and $L_i + 1$. So

$$\left[\frac{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}{\left[\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)\right]^{\frac{1}{a}}}.D^{-l_i\left(\frac{a-1}{a}\right)}\right]^a \leq \left[\frac{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}{\left[\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)\right]^{\frac{1}{a}}}.D^{-L_i\left(\frac{a-1}{a}\right)}\right]^a$$

$$< D\left[\frac{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}{\left[\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)\right]^{\frac{1}{a}}}.D^{-L_i\left(\frac{a-1}{a}\right)}\right]^a \tag{2.3.2}$$

But,

$$\left[\frac{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}{\left[\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)\right]^{\frac{1}{a}}}.D^{-L_i\left(\frac{a-1}{a}\right)}\right]^a = \frac{\sum_{i=1}^n \left(\mu_A^{ab}(x_i) + \left(1-\mu_A(x_i)\right)^{ab}\right)}{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}$$

Now, from (2.3.2)

$$\left[\frac{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}{\left[\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)\right]^{\frac{1}{a}}}.D^{-l_i\left(\frac{a-1}{a}\right)}\right]^a \leq \frac{\sum_{i=1}^n \left(\mu_A^{ab}(x_i) + \left(1-\mu_A(x_i)\right)^{ab}\right)}{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}$$

$$< D.\left[\frac{\sum_{i=1}^n \left(\mu_A^{ab}(x_i) + \left(1-\mu_A(x_i)\right)^{ab}\right)}{\sum_{i=1}^n \left(\mu_A^b(x_i) + \left(1-\mu_A(x_i)\right)^b\right)}\right]$$

i.e.

$$-\left[\frac{\sum_{i=1}^{n}\left(\mu_A^b(x_i)+\left(1-\mu_A(x_i)\right)^b\right)}{\left[\sum_{i=1}^{n}\left(\mu_A^b(x_i)+\left(1-\mu_A(x_i)\right)^b\right)\right]^{\frac{1}{a}}}.D^{-l_i\left(\frac{a-1}{a}\right)}\right]^a \geq -\frac{\sum_{i=1}^{n}\left(\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}\right)}{\sum_{i=1}^{n}\left(\mu_A^b(x_i)+\left(1-\mu_A(x_i)\right)^b\right)}$$

$$> -D.\left[\frac{\sum_{i=1}^{n}\left(\mu_A^{ab}(x_i)+\left(1-\mu_A(x_i)\right)^{ab}\right)}{\sum_{i=1}^{n}\left(\mu_A^b(x_i)+\left(1-\mu_A(x_i)\right)^b\right)}\right]$$

i.e.

$$L_a^b-\frac{1}{a-1}\geq V_a^b(A)-\frac{1}{a-1}>V_a^b(A).D-\frac{D}{a-1}$$

i.e.

$$(a-1)L_a^b\geq(a-1)V_a^b(A)>(a-1)V_a^b(A).D-(D-1).$$

Thus, the required result is obtained.

Conclusion:

In this paper, we prove some coding theorems for discrete noiseless channel using average code word length L_a^b as well as derive bounds in the context of a new generalization of fuzzy information measures.

References:

[1] **Aczel, J. (1975):** On Shannon's inequality optimal coding and characterization of Shannon and Renyi's entropies, Institute Novonal De Alta Mathematics Symposia Maths 15, 153-179.

[2] **Boekee, D. E. and Lubbe, J. C. A. (1980):** The R-norm information measures, Information and Control 45, 136-155.

[3] **Campbell, L. L. (1965):** A coding theorem and Renyi's entropy, Information and Control, Vol. 8, pp. 301-312.

[4] **De Luca, A. and Termini, S. (1972):** A definition of Non-probabilistic Entropy in the of fuzzy set theory, Information and Control 20, 301-312.

[5] **Havrada, J. H. and Charvat, F. (1967):** Quantification methods of classificatory processes, the concepts of structural α entropy, Kybernetika, Vol. 3, pp.30-35.

[6] **Kaufmann, A. (1975):** Introduction to the Theory of Fuzzy Subsets, Vol. 1, Fundamental Theoretical Elements, Academic Press, New York.

[7] **Kraft, L. J. (1949):** A device for quantizing grouping and coding amplitude modulates pulses, M.S. Thesis, Department of Electrical Engineering, MIT, Cambridge.

[8] **Kapur, J. N. (1997):** Measures of Fuzzy Information, Mathematical Science Trust Society, New Delhi.
[9] **Lowen, R. (1996):** Fuzzy Set Theory Basic Concepts, Techniques and Bibliography, Kluwer Academic Publication, Applied Intelligence 31, 283-291.
[10] **Mitter, J. and Mathur, Y. D. (1972):** Comparison of entropies of power distribution, ZAMM, 52, 239-240.
[11] **Nguyen, H. T. and Walker, E. A. (1997):** A first course in fuzzy logic, C.R.C. Press Inc.
[12] **Renyi, A. (1961):** On measure of entropy and information, Proceeding Fourth Berkely Symposium on Mathematical Statistics and Probability, University of California Press 1, 547-561.
[13] **Shannon, C. E. (1948):** A mathematical theory of communication, Bell System Technical Journal 27, 379-423, 623-659.
[14] **Tsallis, C. (1988):** Possible Generalization of Boltzmann-Gibbs Statistics, J. Stat. Phy., 52, 479.
[15] **Verma, R. K. and Verma, Babita (2012):** A New Approach in Mathematical Theory of Communication (A New Entropy with its Application), Lambert Academic Publishing Edn., pp. 83-86.
[16] **Van Der Lubbe, J. C. A. (1978):** On certain coding theorems for the information of order α and type β. In. Trans., 8^{th} Prague conf. Inf. Theory, D. Reidell, Dordrecth 253-256.
[17] **Zadeh, L. A. (1965):** Fuzzy Sets, Information and Control, Vol. 8, pp. 94-102.