A Measure of Divergence between Fuzzy Sets with Advancements in Information Theory

Priti Gupta
Department of Statistics
Maharshi Dayanand University
Rohtak, Haryana, INDIA

H. D. Arora
Department of Mathematics
Amity University, Noida,
U.P., INDIA

Pratiksha Tiwari
Department of CS
D I A S, GGSIP University,
Delhi, INDIA

ABSTRACT
Uncertainty and fuzziness are involved in most of the real world problems. In present communication we have studied various advancements in fuzzy information theory and proposed a new measure of divergence between two fuzzy sets. Further, a measure of fuzzy directed divergence was developed based on the proposed divergence measure. Some properties of this fuzzy directed divergence measure were established.

General Terms
Information Theory, Fuzzy Directed Divergence

Keywords
Information theory, Fuzzy Information theory, Entropy, Divergence, Fuzzy Directed Divergence.

1. INTRODUCTION
Information theory studies all problems related to the entity called communication system. In other words it deals with the study of problems concerning any system that includes information dispensation, storage, retrieval and decision making. Fundamental theorem of information theory states that "It is possible to transmit information over a noisy channel at any rate less than channel capacity with an arbitrary small probability of error". By reducing the transmission rate to zero probability of error can be reduced to zero, but according to fundamental theorem of information theory is that in order to achieve higher reliability their no need to reduce the transmission rate to zero but by reducing it to channel capacity and this can be achieved by means of coding. Information theory considered to be identified by Shannon (1948).

Zadeh gave fuzzy set theory in 1965 . Uncertainty and fuzziness is the basic nature of human thinking and of many real world objectives. Fuzziness is found in our decision, in our language and in the way we process information. Kullback and Liebler (1951) quantified the measure of information associated with the two probability distributions of discrete random variables named as directed divergence. Bhandari and Pal (1993) give a fuzzy information measure for discrimination of a fuzzy set $\tilde{A}$ relative to some other fuzzy set $\widetilde{B}$ called as Fuzzy divergence and it found various applications in real world such as image segmentation, medical sciences, pattern recognition, fuzzy clustering etc.

The present communication proposed a new measure of fuzzy divergence and the corresponding fuzzy directed divergence is defined and some of its properties are studied.

The paper is divided into three sections. Section II contains some preliminaries about measure of uncertainty, directed divergence, fuzzy divergence and fuzzy directed divergence. Section III contains various developments in information theory. Section IV includes new aggregation operators, divergence measure, fuzzy directed divergence and Section V includes conclusion.

2. PRELIMINARIES
2.1 Information Measures
In communication system, the source of messages can be a person or machine that generates the messages, the encoder convert messages in to an object which is suitable for transmission, such as a sequence of binary digits (digital computer applications), channel is a medium over which the coded message is transmitted, decoder convert the received output from the channel and try to convert the received output in to the original message to transport it to the destination. But this cannot be done with absolute consistency due to existence of some disorder in the system which is also termed as noise. We can define information transmitted over channel can be sent at any rate less than channel capacity with an very small probability of error (with the help of a coding system) with the help of Entropy (also termed as measure of uncertainty). It also satisfies four Axioms for Uncertainty Measures:

Axiom 1: $f(M) = H\left(1/_M, 1/_M, \dots\dots.., 1/_M\right)$ is a monotonically increasing function.

Axiom 2:

$$f(ML) = f(M) + f(L), M, L = 1,2,3, \dots\dots….$$

Axiom 3:

$$H(p_1, p_2, \dots\dots. p_M)$$
$$= H(p_1 + \cdots. + p_r, p_{r+1} + \cdots. + p_m)$$
$$+ (p_1 + \cdots. + p_r) H\left(\frac{p_1}{\sum_{i=1}^{r} p_i}, \dots.., \frac{p_r}{\sum_{i=1}^{r} p_i}\right)$$
$$+ (p_{r+1}$$
$$+ \cdots. + p_M) H\left(\frac{p_{r+1}}{\sum_{i=r+1}^{M} p_i}, \dots.., \frac{p_M}{\sum_{i=r+1}^{M} p_i}\right)$$

Axiom 4: $H(p, 1-p)$ is continuous function of p.

The function which satisfies the four axioms of uncertainty is

$$H(p_1, p_2, \dots.., p_M) = -K \sum_{i=1}^{M} p_i log(p_i)$$

where K is any positive number and the logarithm base is a number greater than 1.By taking K=1 and logarithm base 2 , equation (1) reduces to

$H(p_1, p_2, \ldots\ldots, p_M) = -\sum_{i=1}^{M} p_i log(p_i)$, the expression is known a measure of uncertainty or Shannon's Entropy.

Kullback and Liebler (1951) quantified the measure of information associated with the two probability distributions

$$p = (p_1, p_2, \ldots\ldots, p_n) and\ q = (q_1, q_2, \ldots\ldots, q_n)$$

of discrete random

variables , $D(p \parallel q) = \sum_{i=1}^{n} p_i log \frac{p_i}{q_i}$ known as directed divergence. There exist other measures of divergence on set of probabilities, with diverse names discrimination, distance etc.

The natural properties of directed divergence are divergence is non-negative function; it becomes zero when two sets coincide.

Fuzzy set theory gave a whole new facet to set theory which considers either element belongs to or does not belong to a set. A fuzzy set $\tilde{A}$ is subset of universe of discourse X, is defined as $\tilde{A} = \{\langle x, \mu_{\tilde{A}}(x)\rangle / x \in X\}$, where $\mu_{\tilde{A}} : X \to [0,1]$ is a membership function of $\tilde{A}$. The value of $\mu_{\tilde{A}}(x)$ describes the degree of belongingness of $x \in X$ in $\tilde{A}$.

Fuzzy entropy deals with vagueness and ambiguous uncertainties, whereas Shannon entropy deals with probabilistic uncertainties. De Luca and Termini (1972) characterized the fuzzy entropy and introduced a set of following properties (1–4) for which fuzzy entropy should satisfy them:

1. Fuzzy entropy is minimum iff set is crisp.
2. Fuzzy entropy is maximum when membership value is 0.5.
3. Fuzzy entropy decreases if set is sharpened.
4. Fuzzy entropy of a set is same as its complement.

Bhandari and Pal (1993) gave a range of measures of fuzzy entropy and measure of fuzzy divergence corresponding to fuzzy sets $\tilde{A}$ and $\tilde{B}$. Let Ω be a Universal set and F(Ω) be the collection of fuzzy subsets of Ω. A mapping D: F(Ω)×F(Ω)→R is called divergence between two fuzzy subsets if it satisfies following properties for any$\tilde{A}, \tilde{B}, \tilde{C} \in F(\Omega)$:

1. $D(\tilde{A}, \tilde{B})$is non-negative.
2. $D(\tilde{A}, \tilde{B})$= $D(\tilde{B}, \tilde{A})$
3. $D(\tilde{A}, \tilde{B}) = 0$if $\tilde{A} = \tilde{B}$
4. Max{ $D(\tilde{A} \cup \tilde{C}, \tilde{B} \cup \tilde{C}), D(\tilde{A} \cap \tilde{C}, \tilde{B} \cap \tilde{C})\} \leq D(\tilde{A}, \tilde{B})$

The fuzzy directed divergence $D(\tilde{A}, \tilde{B}) = \sum_{i=1}^{n} \left[\mu_A(x_i) log \frac{\mu_A(x_i)}{\mu_B(x_i)} + (1 - \mu_A(x_i)) log \frac{(1-\mu_A(x_i))}{(1-\mu_B(x_i))}\right]$ is given by Bhandari and Pal(1993), where

$\mu_A(x_1), \mu_A(x_2), \ldots\ldots\ldots, \mu_A(x_n)$ are the degree of belongingness of $x_i \in X$ in $\tilde{A}$ and $\mu_B(x_1), \mu_B(x_2), \ldots\ldots\ldots, \mu_B(x_n)$ are the degree of belongingness of $x_i \in X$ in $\tilde{B}$ respectively.

Operations on fuzzy sets are termed as aggregation operators such as fuzzy union and fuzzy intersection. These operations are used to combine two or more fuzzy sets into one. An aggregation operation (Klir and Folger, 1988) defined as a function $A: [0,1]^n \to [0,1]$ satisfying:

A(0,0,…..0)=0 and A(1,1,…..,1)=1

A is monotonic in each argument.

2.2 Recent Developments

This section includes various developments in this area currently. Bhatia and Singh (2012), proposes a measure of arithmetic–geometric directed divergence of two arbitrary fuzzy sets A and B is as

$$T(A, B) = \sum_{i=1}^{n} \left[\frac{\mu_A(x_i) + \mu_B(x_i)}{2} log \frac{\mu_A(x_i) + \mu_B(x_i)}{2\sqrt{\mu_A(x_i)\,\mu_B(x_i)}} + \frac{2-\mu_A(x_i) - \mu_B(x_i)}{2} log \frac{2 - \mu_A(x_i) - \mu_B(x_i)}{2\sqrt{(1-\mu_A(x_i))(1-\mu_B(x_i))}}\right]$$

and defined generalized triangular discrimination between two arbitrary fuzzy sets A and B as follows:

$$\Delta_\alpha(A, B) = \sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^{2\alpha} \left[\frac{1}{\left(\mu_A(x_i) + \mu_B(x_i)\right)^{2\alpha-1}} + \frac{1}{\left(2 - \mu_A(x_i) - \mu_B(x_i)\right)^{2\alpha-1}}\right]$$

They also defined a new (α, β) class of measure of fuzzy directed divergence for two arbitrary sets A and B as

$$D_\alpha^\beta(A, B) = \frac{1}{\beta-1} \sum_{i=1}^{n} \left[\left(\mu_A(x_i)^\alpha \mu_B(x_i)^{1-\beta} + \left(1 - \mu_A(x_i)\right)^{1-\alpha} (1 - \mu_B(x_i))^{1-\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right], \alpha > 0, \alpha \neq 1, \beta > 0, \beta \neq 1$$

and introduced (α, β) generalized arithmetic-geometric measure of fuzzy directed divergence $T_\alpha^\beta(A, B) = \frac{1}{2}\left[D_\alpha^\beta\left(\frac{A+B}{2}, A\right) + D_\alpha^\beta\left(\frac{A+B}{2}, B\right)\right]$

and $T_\alpha^\beta(A, B) = T(A, B) at\ \alpha = \beta = 1$.

A new measure of fuzzy directed divergence for two Fuzzy sets A and B,

$$M_{H^*\ \ A^*}^F(A, B) = \sum_{i=1}^{n} \frac{\left(\mu_A(x_i) - \mu_B(x_i)\right)^2}{2} \left[\frac{1}{\mu_A(x_i) - \mu_B(x_i)} + \frac{1}{2-\mu_A(x_i) - \mu_B(x_i)}\right]$$

, Where, $A^*: [0,1]^2 \to [0,1]$ such that

$A^*(a,b)=\frac{a+b}{2}$ and $H^*:[0,1]^2\to[0,1]$ such that $H^*(a,b)=\frac{a^2+b^2}{a+b}$ was defined by Bhatia and Singh(2013(a)), they also discussed application of new directed divergence measure in images segmentation.

Bhatia and Singh (2013(b)), introduced three new divergence measures between fuzzy sets

$$M_1(A,B)=\sum_{i=1}^{n}\frac{\left(\mu_A(x_i)-\mu_B(x_i)\right)^2}{2}\left[\frac{1}{\mu_A(x_i)+\mu_B(x_i)}+\frac{1}{2-\mu_A(x_i)-\mu_B(x_i)}\right]$$

$$M_2(A,B)=\sum_{i=1}^{n}\left[\frac{\left(\mu_A(x_i)\right)^2+\left(\mu_B(x_i)\right)^2}{\mu_A(x_i)+\mu_B(x_i)}-\sqrt{\mu_A(x_i)\mu_B(x_i)}+\frac{\left(1-\mu_A(x_i)\right)^2+\left(1-\mu_B(x_i)\right)^2}{2-\mu_A(x_i)-\mu_B(x_i)}-\sqrt{\left(1-\mu_A(x_i)\right)\left(1-\mu_B(x_i)\right)}\right]$$

$$M_3(A,B)=\sum_{i=1}^{n}\left(\mu_A(x_i)-\mu_B(x_i)\right)^2\left[\frac{1}{\mu_A(x_i)+\mu_B(x_i)}+\frac{1}{2-\mu_A(x_i)-\mu_B(x_i)}\right]$$

and some properties of these divergence measures. They also defined three aggregation functions corresponding to divergence measures given above as $\Phi_1(\vec{p})=\frac{1}{n}\sum_{i=1}^{n}\frac{p_i}{2-p_i}$, $\Phi_2(\vec{p})=\frac{1}{2n}\sum_{i=1}^{n}p_i+\frac{1+(1-p_i)^2}{2-p_i}-\sqrt{1-p_i}$ and $\Phi_3(\vec{p})=\frac{2}{n}\sum_{i=1}^{n}\frac{p_i}{2-p_i}$ respectively.

Verma, Dewangan, Jha(2012), defined a measure of entropy as $V_a(P)=\sum_{i=1}^{n}\ln(1+ap_i)-\sum_{i=1}^{n}lnp_i-\ln(1+a)$, $a>0$ for probability distribution, $P=(p_1,p_2,\ldots\ldots,p_n)$ and its corresponding measure of directed divergence is defined as $D_a(P:Q)=\sum_{i=1}^{n}q_i ln\frac{p_i}{q_i}-\sum_{i=1}^{n}q_i\ln\left(\frac{q_i+ap_i}{q_i}\right)+\ln(1+a)$, $a>0$ and corresponding measure of fuzzy directed divergence is

$$D(\mathrm{A,B})=\sum_{i=1}^{n}\mu_B(x_i)ln\left(\frac{\mu_A(x_i)}{a\mu_A(x_i)+\mu_B(x_i)}\right)+\sum_{i=1}^{n}\left(1-\mu_B(x_i)\right)ln\left(\frac{1-\mu_A(x_i)}{1+a-a\mu_A(x_i)-\mu_B(x_i)}\right)+\ln(1+a),\quad a>0$$

and their properties were studied.

3. FUZZY DIRECTED DIVERGENCE

In this section we have defined as two new binary aggregation operators and corresponding to these operators a new divergence measure has been introduced. Further, new directed divergence measure for two fuzzy sets have been defined.

Aggregation operators are defined as $G:[0,1]^2\to[0,1]$, $G(a,b)=a^{3/2}b^{3/2}$ and $K:[0,1]^2\to[0,1]$, $K(a,b)=\frac{a^4+b^4}{a+b}$, both the functions clearly satisfies all the properties of aggregation operators. Let us define a new divergence as

$$D_{G,K}(P,Q)=\sum_{i=1}^{n}\left[\frac{p_i^4+q_i^4}{p_i+q_i}-p_i^{3/2}q_i^{3/2}\right] \quad \ldots\ldots\ldots\ldots.. (1)$$

Where $P=(p_1,p_2,\ldots\ldots,p_n)\ and\ Q=(q_1,q_2,\ldots\ldots.,q_n)$

Equation (1) is clearly non-negative

As $\frac{p_i^4+q_i^4}{p_i+q_i}-p_i^{3/2}q_i^{3/2}=\frac{p_i^4+q_i^4-p_i^{3/2}q_i^{3/2}(p_i+q_i)}{p_i+q_i}=\frac{p_i^4+q_i^4-p_i^{5/2}q_i^{3/2}-p_i^{3/2}q_i^{5/2}}{p_i+q_i}$

$=\frac{\left(p_i^4-p_i^{3/2}q_i^{5/2}\right)+(q_i^4-p_i^{5/2}q_i^{3/2})}{p_i+q_i}=\frac{p_i^{3/2}\left(p_i^{5/2}-q_i^{5/2}\right)-q_i^{3/2}(p_i^{5/2}-q_i^{5/2})}{p_i+q_i}$

$=\frac{(p_i^{3/2}-q_i^{3/2})\left(p_i^{5/2}-q_i^{5/2}\right)}{p_i+q_i}\geq 0\ for\ each\ i=1,2,\ldots\ldots\ldots,n$, because both $(p_i^{3/2}-q_i^{3/2})$ and $\left(p_i^{5/2}-q_i^{5/2}\right)$ are simultaneously negative or positive.

Hence $D_{G,K}(P,Q)$ is a non-negative function with minimum value $D_{G,K}(P,Q)=0\ as\ P=Q$ and hence a valid measure of divergence. Hence equation (1) is a valid measure of divergence.

Let us compare the membership values by comparing the fuzziness of A, B with the fuzziness of the intermediate subset. Consider an example let us consider the universe consisting of four elements A, B, C and D with membership values given in the following Table 1:

Table 1: Membership Values

	x_1	x_2	x_3	x_4
A	0.25	0.8	0.5	0.3
B	0.2	0.7	0.5	0.2
C	0.04	0.5	0.7	0.5
D	0.89	0.01	0.1	0.99

For these sets we

obtain $D_{G,K}(A,B) = 0.019827$, $D_{G,K}(A,C) = 0.197909$ and $D_{G,K}(A,D) = 1.636029.$ The divergence measure shows that A is quite similar to B and different form D.

Measure of fuzzy directed divergence between two fuzzy sets A and B with respect to the proposed divergence in equation (1) is defined as

$$F_{G,K}(A,B) = \sum_{i=1}^{n}\left[\frac{\mu_A(x_i)^4 + \mu_B(x_i)^4}{\mu_A(x_i) + \mu_B(x_i)} - \mu_A(x_i)^{3/2}\mu_B(x_i)^{3/2} + \frac{(1-\mu_A(x_i))^4 + (1-\mu_B(x_i))^4}{2-\mu_A(x_i)-\mu_B(x_i)} - (1-\mu_A(x_i))^{3/2}(1-\mu_B(x_i))^{3/2}\right]$$

Clearly, $F_{G,K}(A,B)$ is non-negative, $F_{G,K}(A,B) = F_{G,K}(B,A)$ and $F_{G,K}(A,A) = 0$. To prove the fourth property , divide the universe of discourse (Ω) in to following six subsets as

$$\Omega = \{x \in X/\mu_A(x) \le \mu_B(x) \le \mu_C(x)\} \cup \{x \in X/\mu_A(x) \le \mu_C(x) < \mu_B(x)\} \cup \{x \in X/\mu_B(x) < \mu_A(x) \le \mu_C(x)\} \cup \{x \in X/\mu_B(x) \le \mu_C(x) < \mu_A(x)\} \cup \{x \in X/\mu_C(x) < \mu_A(x) \le \mu_B(x)\} \cup \{x \in X/\mu_C(x) < \mu_B(x) < \mu_A(x)\}$$

which we will denote as $\Omega_1, \Omega_2, \Omega_3, \Omega_4, \Omega_5, \Omega_6$ respectively. We compute $F_{G,K}(A,B)$ in each of the subsets Ω_i , $i = 1,2,3,4,5,6$. Then combine the results, thus we obtain $F_{G,K}(A,B)$ for the universe of discourse.

In

Ω_1, $A \cup C \Leftrightarrow \mu_{A\cup C}(x) = \max\{\mu_A(x), \mu_C(x)\} = \mu_C(x)$ $B \cup C \Leftrightarrow \mu_{B\cup C}(x) = \max\{\mu_B(x), \mu_C(x)\} = \mu_C(x)$

$$A \cap C \Leftrightarrow \mu_{A\cap C}(x) = \min\{\mu_A(x), \mu_C(x)\} = \mu_A(x)$$

$$B \cap C \Leftrightarrow \mu_{B\cap C}(x) = \min\{\mu_B(x), \mu_C(x)\} = \mu_B(x)$$

Therefore, $F_{G,K}(A \cup C, B \cup C) = 0$ and $F_{G,K}(A \cap C, B \cap C) = F_{G,K}(A,B)$

Hence property four, max { $F_{G,K}(A \cup C, B \cup C)$, $F_{G,K}(A \cap C, B \cap C)$} ≤ $F_{G,K}(A,B)$ holds good for Ω_1, similarly the inequality holds for $\Omega_2, \Omega_3, \Omega_4, \Omega_5$ *and* Ω_6. Thus $F_{G,K}(A,B)$ is a valid measure of fuzzy directed divergence.

Theorem: Let A and B two fuzzy sets then following properties can be verified for $F_{G,K}(A,B)$:

1. $F_{G,K}(A \cup B, A \cap B) = F_{G,K}(A,B)$.
2. $F_{G,K}(A,\bar{A}) = 2n$, when A is a crisp set i.e. $\mu_A(x) = 0 \text{ or } 1$
3. $F_{G,K}(A,B) = F_{G,K}(\bar{A}, \bar{B})$

Proof: Divide the universe of discourse into tow subsets as $\Omega = \{x \in X/\mu_A(x) \le \mu_B(x)\} \cup \{x \in X/\ \mu_B(x) < \ \mu_A(x)\}$, which we will denote as Ω_1, Ω_2 respectively.

In

$\Omega_1, A \cup B \Leftrightarrow \mu_{A\cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} = \mu_B(x)$ And $A \cap B \Leftrightarrow \mu_{A\cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} = \mu_A(x)$ $F_{G,K}(A \cup B, A \cap B) = F_{G,K}(B,A) = F_{G,K}(A,B)$ Similarly we can prove the result for Ω_2. Hence property 1 holds. Cleary property 2 and 3 holds.

4. CONCLUSION

In this paper we have presented various information theoretic advancements in the area and a new measure of fuzzy directed divergence has been defined using aggregation operators and some of its mathematical properties have been proved.

5. REFERENCES

[1] Bhandari, D. and Pal, N. R., 1993. Some new information measures for fuzzy sets, Information Sciences, Vol. 67, pp. 204 - 228.

[2] Bhatia P.K. and Singh S., 2013(a). A New Measure of Fuzzy Directed Divergence and Its Application in Image Segmentation, I.J. Intelligent Systems and Applications, Vol.4, pp.81-89.

[3] Bhatia P.K. and Singh S., 2013(b). On Some Divergence Measures between Fuzzy Sets and Aggregation Operations, AMO-Advanced Modeling and Optimization, Vol.15, No. 2,pp 235-248.

[4] Bhatia P.K. and Singh S., 2012. Three Families of Generalized Fuzzy Directed divergence, AMO-Advanced Modeling and Optimization, Volume 14, Number 3,pp.599-614.

[5] De Luca A. and Termini S., 1972. A definition of a Non-probabilistic Entropy in Setting of Fuzzy Sets, Information and Control, Vol. 20, pp.301-312.

[6] Klir G.J. and Folger T.A., Fuzzy Sets, Uncertainty, and Information, Prentice-Hall International (UK) Limited, 1988.

[7] Kullback, S. and Leibler, R.A., 1952. On Information and Sufficiency, Ann. Math. Stat., Vol 22, pp. 79 - 86.

[8] Montes S., Couso I., Gil P. and Bertoluzza C., 2002. Divergence Measure between Fuzzy Sets, International Journal of Approximate Reasoning, Vol. 30,pp. 91-105.

[9] Shannon C.E., 1948. A Mathematical Theory of Communication, The Bell System Technical Journal, Vol. 27, July, pp.623-656.

[10] Verma R.K, Dewangan C.L., Jha P., 2012. An Unorthodox Parametric Measures of Information and Corresponding Measures of Fuzzy Information, International Journal of Pure and Applied Mathematics, Vol. 76, No. 4 , pp.599-614.

[11] Zadeh, L.A., 1965. Fuzzy Sets, Information and Control, Vol 8, pp. 338- 353.