LETTER

Study on Entropy and Similarity Measure for Fuzzy Set

Sang-Hyuk LEE[†], ***Member***, **Keun Ho RYU**[††], ***Nonmember***, ***and*** **Gyoyong SOHN**[††a)], ***Member***

SUMMARY In this study, we investigated the relationship between similarity measures and entropy for fuzzy sets. First, we developed fuzzy entropy by using the distance measure for fuzzy sets. We pointed out that the distance between the fuzzy set and the corresponding crisp set equals fuzzy entropy. We also found that the sum of the similarity measure and the entropy between the fuzzy set and the corresponding crisp set constitutes the total information in the fuzzy set. Finally, we derived a similarity measure from entropy and showed by a simple example that the maximum similarity measure can be obtained using a minimum entropy formulation.
key words: *similarity measure, distance measure, fuzzy entropy*

1. Introduction

Average values and standard deviation analyze statistical information from different points of view when a heuristic approach is adopted, and hence, they are sometimes inconsistent with each other. To analyze ambiguous data, we have to consider a data set as a fuzzy set with a degree of membership. Entropy and similarity analyses are essential for studying the total data information of fuzzy sets. The characterization and quantification of fuzziness are important and affect the management of uncertainty in the modeling and design of many systems. The fact that the entropy of a fuzzy set is a measure of its fuzziness has been established by previous researchers [1]–[6]. Zadeh was the first to propose fuzzy entropy as a measure of fuzziness; Pal and Pal analyzed classical Shannon information entropy; Kosko considered the relationship between distance measure and fuzzy entropy; Liu proposed axiomatic definitions of entropy, distance measures, and similarity measures and discussed the relationships among these three concepts. Bhandari and Pal presented a measure of fuzzy information for distinguishing between fuzzy sets. Further, Ghosh used fuzzy entropy in neural networks.

The degree of similarity between two or more data sets has a central role in decision making, pattern classification, etc., [7]–[10]. Thus far, numerous researchers have carried out research on deriving similarity measures [11]–[14]. Methods based on fuzzy numbers enable the simple derivation of similarity measures. However, derived similarity measures are restricted to triangular or trapezoidal membership functions [11], [12]. In contrast, similarity measures based on the distance measure are applicable to general fuzzy membership functions, including nonconvex fuzzy membership functions [13], [14].

The correlation between entropy and similarity for fuzzy sets has been presented from various viewpoints [15]. Liu also proposed a relation between distance and similarity measures; in his paper, the sum of distance and similarity constitutes the total information [4]. In this paper, we analyze the relationship between the entropy and similarity measures for fuzzy sets and the corresponding numeric data sets. We derive fuzzy entropy and similarity measures by using the distance measure. We verify the total information property, which combines the similarity and entropy measures. The fuzzy entropy between two comparative data sets enables us to also obtain the similarity measure by using the total information property.

In the following section, we discuss the relationship between entropy and similarity for a fuzzy set. We obtain the corresponding crisp set for a fuzzy set satisfying minimum fuzzy entropy. We also discuss the previously obtained fuzzy entropy and similarity measure. In Sect. 3, we derive the procedure for obtaining the similarity measure from the fuzzy entropy. Furthermore, the maximum similarity for a fuzzy set is obtained in a simple example. The conclusions are stated in Sect. 4.

2. Relation between Entropy and Similarity Measure

Data uncertainties are inherent to fuzzy sets with membership functions. We have proposed that fuzzy entropy be used to measure the uncertainties [13]. We require two comparative sets to develop fuzzy entropy. One is a fuzzy set and the other is the corresponding crisp set. On the basis of the definition of fuzzy entropy, numerous fuzzy entropies are presented. The fuzzy membership function pair is illustrated in Fig. 1. An analysis of the entropy for fuzzy set shows that it is essential to consider the corresponding crisp set. A_{near} represents the crisp set "near" fuzzy set A. The value of $A_{0.5}$ is one when $\mu_A(x) \geq 0.5$ and is zero otherwise. A_{far} is the complement of A_{near}, i.e., $A^C_{near} = A_{far}$.

We proposed the fuzzy entropy of fuzzy set A with respect to A_{near} as follows [13].

$$e(A, A_{near}) = d(A \cap A_{near}[1]_X) + d(A \cup A_{near}[0]_X) - 1 \quad (1)$$

Manuscript received February 12, 2009.
Manuscript revised May 21, 2009.
[†]The author is with School of Mechatronics, Changwon National University #9 sarim-dong, Changwon, Geyeongam 641–773, Korea.
[††]The authors are with the Department of Computer Science, Chungbuk National University #81 Gaesin-dong, Cheongju, Republic of Korea.
a) E-mail: gysohn@dblab.chungbuk.ac.kr
DOI: 10.1587/transinf.E92.D.1783

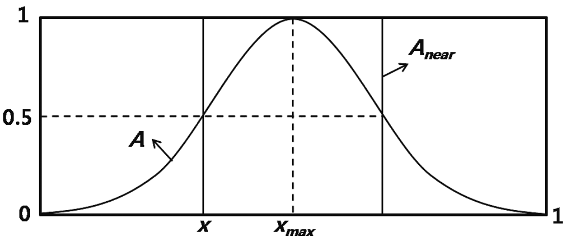

Fig. 1 Fuzzy membership function pairs of A and $A_{near} = A_{0.5}$.

where $d(A \cap A_{near}[1]_X) = \frac{1}{n}\sum_{i=1}^{n} |\mu_{A\cap A_{near}}(x_i) - 1|$ is satisfied. $[0]_X$ and $[1]_X$ are the fuzzy sets in which the value of the membership functions are zero and one, respectively, for the universe of discourse; d also satisfies the Hamming distance measure.

Equation (1) does not give the normal fuzzy entropy. The normal fuzzy entropy can be obtained by multiplying the right-hand side of Eq. (1) by two, which satisfies maximal fuzzy entropy is one. The fuzzy entropy in Eq. (1) satisfies for all value of crisp set A_{near}. Hence, $A_{0.1}$ and $A_{0.5}$ or some other $A_{0.X}$ can be satisfied. Now, it is interesting to search for the value of $A_{0.X}$ that is a maximum or minimum.

Equation (1) is rewritten as follows:

$$e(A, A_{near}) = 2\int_0^x \mu_A(x)dx + 2\int_x^{x_{\max}} 1 - \mu_A(x)dx. \quad (2)$$

Let $\frac{d}{dx}M_A(x) = \mu_A(x)$; $e(A, A_{near})$ has been shown to be

$$e(A, A_{near}) = 2M_A(x)|_0^x + 2(x_{\max} - x) - 2M_A(x)|_x^{x_{\max}}.$$

The maxima or minima are obtained by differentiation:

$$\frac{d}{dx}e(A, A_{near}) = 2\mu_A(x) - 2 + 2\mu_A(x).$$

Hence, it is clear that the point x satisfying $\frac{d}{dx}e(A, A_{near}) = 0$ is the critical point for the crisp set. This point is given by $\mu_A(x) = 1/2$, i.e., $A_{near} = A_{0.5}$.

The fuzzy entropy between A and $A_{0.5}$ has a minimum value because $e(A)$ attains maxima when the corresponding crisp sets are $A_{0.0}$ and $A_{x_{\max}}$. Hence, for a nonconvex and symmetric fuzzy set, the minimum entropy of the fuzzy set is equal to that of the crisp set $A_{0.5}$. This indicates that the corresponding crisp set that has the least uncertainty or the greatest similarity with the fuzzy set is $A_{0.5}$.

All the studies on similarity measures deal with derivations of similarity measures and applications in the distance-measure-based computation of the degree of similarity. Liu has also proposed an axiomatic definition of the similarity measure [4]. The similarity measure $\forall A, B \in F(X)$ and $\forall D \in P(X)$ has the following four properties:

(S1)$s(A, B) = s(B, A),\ \forall A, B \in F(X)$

(S2)$s(D, D^c) = 0,\ \forall D \in P(X)$

(S3)$s(C, C) = \max_{A,B\in F} s(A, B),\ \forall C \in F(X)$

(S4)$\forall A, B, C \in F(X)$, if $A \subset B \subset C$, then $s(A, B) \geq s(A, C)$ and $s(B, C) \geq s(A, C)$

where $F(X)$ denotes a fuzzy set, and $P(X)$ is a crisp set.

The proposed similarity measure between A and A_{near} is presented in Theorem 2.1. We verify the usefulness of this measure through a proof of this theorem.

Theorem 2.1 $\forall A \in F(X)$ and the crisp set A_{near} in Fig. 1,

$$s(A, A_{near}) = d(A \cap A_{near}, [0]_X) + d(A \cup A_{near}, [1]_X) \quad (3)$$

is a similarity measure.

Proof. (S1) follows from Eq. (3), and for crisp set D, it is clear that $s(D, D^C) = 0$. Hence, (S2) is satisfied. (S3) indicates that the similarity measure of two identical fuzzy sets $s(C, C)$ attains the maximum value among various similarity measures with different fuzzy sets A and B since $d(C\cap C, [0]_X)+d(C\cup C, [1]_X)$ represents the entire region in Fig. 1. Finally, from $d(A\cap A_{1near}, [0]_X) \geq d(A\cap A_{2near}, [0]_X)$ and $d(A \cup A_{1near}, [1]_X) \geq d(A \cup A_{2near}, [1]_X)$, $A \subset A_{1near} \subset A_{2near}$; it follows that

$$s(A, A_{1near}) = d(A\cap A_{1near}, [0]_X)+d(A\cup A_{1near}, [1]_X)$$
$$\geq d(A\cap A_{2near}, [0]_X)+d(A\cup A_{2near}, [1]_X) = s(A, A_{2near}).$$

Similarly, $s(A_{1near}, A_{2near}) \geq s(A, A_{2near})$ is satisfied by the inclusion properties $d(A_{1near} \cap A_{2near}, [0]_X) \geq d(A \cap A_{2near}, [0]_X)$ and $d(A_{1near}\cup A_{2near}, [1]_X) \geq d(A\cup A_{2near}, [1]_X)$. ■

The similarity in Eq. (3) represents the areas shared by two membership functions. In our previous studies, we have proposed other similarity measures between two arbitrary fuzzy sets as follows [13], [14]:

For any two sets $A, B \in F(X)$,

$$s(A, B) = 1 - d(A \cap B^C, [0]_X) - d(A \cup B^C, [1]_X) \quad (4)$$

and

$$s(A, B) = 2 - d(A \cap B, [1]_X) - d(A \cup B, [0]_X) \quad (5)$$

are similarity measures between set A and set B.

In Eqs. (4) and (5), fuzzy set B can be replaced by A_{near}. In addition to those in Eqs. (4) and (5), numerous similarity measures that satisfy the definition of a similarity measure can be derived. From Fig. 1, the relationship between data similarity and entropy for fuzzy set A with respect to A_{near} can be determined on the basis of the total area. The total area is one (universe of discourse × maximum membership value = $1\times 1 = 1$); it represents the total amount of information. Hence, the total information comprises the similarity measure and entropy measure, as shown in the following equation:

$$s(A, A_{near}) + e(A, A_{near}) = 1 \quad (6)$$

With the similarity measure in Eq. (5) and the total information expression in Eq. (6), we obtain the following proposition:

Proposition 2.1 In Eq. (6), $e(A, A_{near})$ follows from the similarity measure in Eq. (5):

$$e(A, A_{near}) = 1 - s(A, A_{near})$$
$$= d(A \cap A_{near}, [1]_X) + d(A \cup A_{near}, [0]_X) - 1$$

The above fuzzy entropy is identical to that in Eq. (1). The property given by Eq. (6) is also formulated as follows:

Theorem 2.2 The total information about fuzzy set A and the corresponding crisp set A_{near},

$$s(A, A_{near}) + e(A, A_{near})$$
$$= d(A \cap A_{near}, [0]_X) + d(A \cup A_{near}, [1]_X)$$
$$+ d(A \cap A_{near}, [1]_X) + d(A \cup A_{near}, [0]_X) - 1, \quad (7)$$

equals one.

Proof. Eq. (7) implies that the sum of the similarity measure and fuzzy entropy equals one, which is the total area in Fig. 1. In Eq. (7),

$$d(A \cap A_{near}, [0]_X) + d(A \cap A_{near}, [1]_X) = 1 \text{ and}$$
$$d(A \cup A_{near}, [1]_X) + d(A \cup A_{near}, [0]_X) = 1.$$

Hence, $s(A, A_{near}) + e(A, A_{near}) = 1 + 1 - 1 = 1$ is satisfied. ■
Now, it is clear that the total information about fuzzy set A comprises similarity and entropy measures with respect to the corresponding crisp set.

3. Similarity Measure Design through Entropy

We obtain a similarity measure using fuzzy entropy different from that in Eq. (1). The proposed fuzzy entropy is developed by using the Hamming distances between a fuzzy set and the corresponding crisp set. The following result clearly follows from Fig. 1:

$$e(A, A_{near}) = d(A, A \cap A_{near}) + d(A_{near}, A \cap A_{near}) \quad (8)$$

Eq. (8) represents the difference between A and the corresponding crisp set A_{near}. From Theorem 2.2, the following similarity measure that satisfies Eq. (6) follows:

$$s(A, A_{near}) = 1 - d(A, A \cap A_{near}) - d(A_{near}, A \cap A_{near}) \quad (9)$$

Here, it is interesting to determine whether Eq. (9) satisfies the conditions for a similarity measure.

Proof. (S1) follows from Eq. (9). Furthermore, $s(D, D^C) = 1 - d(D, D \cap D^C) - d(D^C, D \cap D^C)$ is zero because $d(D, D \cap D^C) + d(D^C, D \cap D^C)$ satisfies $d(D, [0]_X) + d(D^C, [0]_X) = 1$. Hence, (S2) is satisfied. (S3) is also satisfied since $d(C, C \cap C) + d(C, C \cap C) = 0$; hence, it follows that $s(C, C)$ is a maximum. Finally,

$$1 - d(A, A \cap B) - d(B, A \cap B) \geq$$
$$1 - d(A, A \cap C) - d(C, A \cap C)$$

because $d(A, A \cap B) = d(A, A \cap C)$ and $d(B, A \cap B) \leq d(C, A \cap C)$ are satisfied for $A \subset B \subset C$. The inequality $s(B, C) \geq s(A, C)$ is also satisfied in a similar manner. ■
Now, by using Eq. (9), we obtain the maximum similarity measure for the fuzzy set. In our previous result, the minimum fuzzy entropy could be obtained when we considered the entropy between the fuzzy sets A and $A_{0.5}$. Hence, it is obvious that the obtained similarity

$$s(A, A_{0.5}) = 1 - d(A, A \cap A_{0.5}) - d(A_{0.5}, A \cap A_{0.5}) \quad (10)$$

represents the maximum similarity measure.

Let us consider the next fuzzy set with membership function $A = \{x, \mu_A(x)\}$:

$$\{(0, 0), (0.1, 0.2), (0.2, 0.4), (0.3, 0.7), (0.4, 0.9), (0.5, 1),$$
$$(0.6, 0.9), (0.7, 0.7), (0.8, 0.4), (0.9, 0.2), (1, 0)\}.$$

The fuzzy entropy and similarity measures calculated using Eqs. (8) and (9) are given in Table 1.

Table 1 Similarity measure between fuzzy set and corresponding crisp set.

Similarity measure	Measure value	Fuzzy entropy	Entropy value
$s(A, A_{0.1})$	0.64	$e(A, A_{0.1})$	0.36
$s(A, A_{0.3})$	0.76	$e(A, A_{0.3})$	0.24
$s(A, A_{0.5})$	0.80	$e(A, A_{0.5})$	0.20
$s(A, A_{0.8})$	0.72	$e(A, A_{0.8})$	0.28
$s(A, A_{0.95})$	0.56	$e(A, A_{0.95})$	0.44

The similarity measure for $s(A, A_{0.5})$ is calculated by using the following equation:

$$s(A, A_{0.5}) = 1 - 1/10(0.2 + 0.4 + 0.4 + 0.2)$$
$$-1/10(0.3 + 0.1 + 0.1 + 0.3) = 0.8.$$

The remaining similarity measures are calculated in a similar manner.

4. Conclusions

Analysis methods for entropy and similarity in fuzzy sets were studied. Fuzzy entropies for fuzzy sets were developed by considering the crisp set "near" the fuzzy set. The minimum entropy can be obtained when the crisp set satisfies $A_{near} = A_{0.5}$. The similarity measure between the fuzzy set and the corresponding crisp set was also derived using the distance measure. Furthermore, we have verified the property that sum of fuzzy entropy and similarity measure between fuzzy set and corresponding crisp set is equal to a constant value. We have also derived the similarity measure using fuzzy entropy and illustrated the calculation of the maximum similarity between the fuzzy set and the corresponding crisp set by presenting a simple example.

Acknowledgments

The authors of this paper were partly supported by the Second Stage of Brain Korea21 Projects and by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (No.R11-2008-014-02002-0).

References

[1] L.A. Zadeh, “Fuzzy sets and systems,” in System Theory, ed. Fox J, pp.29–39, Polytechnic Press, Brooklyn, NY, 1965.
[2] N.R. Pal and S.K. Pal, “Object-background segmentation using new definitions of entropy,” Proc. IEEE, vol.136, no.4, pp.284–295, 1989.
[3] B. Kosko, Neural Networks and Fuzzy Systems, Prentice-Hall, Englewood Cliffs, NJ, 1992.
[4] L. Xuecheng, “Entropy, distance measure and similarity measure of fuzzy sets and their relations,” Fuzzy Sets Syst., vol.52, pp.305–318, 1992.
[5] D. Bhandari and N.R. Pal, “Some new information measure of fuzzy sets,” Inf. Sci., vol.67, no.3, pp.209–228, 1993.
[6] A. Ghosh, “Use of fuzziness measure in layered networks for object extraction: A generalization,” Fuzzy Sets Syst., vol.72, pp.331–348, 1995.
[7] Y. Rébillé, “Decision making over necessity measures through the Choquet integral criterion,” Fuzzy Sets Syst., vol.157, no.23, pp.3025–3039, 2006.
[8] V. Sugumaran, G.R. Sabareesh, and K.I. Ramachandran, “Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine,” Expert Systems with Applications, vol.34, no.4, pp.3090–3098, 2008.
[9] W.S. Kang and J.Y. Choi, “Domain density description for multi-class pattern classification with reduced computational load,” Pattern Recognit., vol.41, no.6, pp.1997–2009, 2008.
[10] F.Y. Shih and K. Zhang, “A distance-based separator representation for pattern classification,” Image Vis. Comput., vol.26, no.5, pp.667–672, 2008.
[11] C.H. Hsieh and S.H. Chen, “Similarity of generalized fuzzy numbers with graded mean integration representation,” Proc. 8th Int. Fuzzy Systems Association World Congr., vol.2, pp.551–555, 1999.
[12] S.J. Chen and S.M. Chen, “Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers,” IEEE Trans. Fuzzy Syst., vol.11, no.1, pp.45–56, 2003.
[13] S.H. Lee, Y.T. Kim, S.P. Cheon, and S.S. Kim, “Reliable data selection with fuzzy entropy,” LNAI, vol.3613, pp.203–212, 2005.
[14] S.H. Lee, J.M. Kim, and Y.K. Choi, “Similarity measure construction using fuzzy entropy and distance measure,” LNAI, vol.4114, pp.952–958, 2006.
[15] S.K. Lin, “Gibbs paradox and the concepts of information, symmetry, similarity and their relationship,” Entropy, vol.10, pp.1–5, 2008.