Chapter 2
Novel Correlation and Entropy Measures of Hesitant Fuzzy Sets

Correlation is one of the most widely used indices in data analysis, pattern recognition, machine learning, decision making, etc. It measures how well two variables move together in a linear fashion. The correlation coefficient, which was originally appeared in Karl Pearson's proposal related to statistics, has been extended into different fuzzy circumstances. Different forms of fuzzy correlation coefficients have been proposed, such as the fuzzy correlation coefficients, the intuitionistic fuzzy correlation coefficients, and the hesitant fuzzy correlation coefficients. Xu and Xia (2011b) defined several correlation coefficients for HFEs. Afterwards, Chen et al. (2013a) proposed a formula to calculate the correlation coefficient between two HFSs. In this chapter, we first point out the weaknesses of the existing correlation coefficients between HFSs, and then introduce some novel correlation coefficient formulas for HFSs. Some new concepts, such as the mean of a HFS, the variance of a HFS and the correlation between two HFSs are defined. Based on these concepts, a novel correlation coefficient formula between two HFSs is introduced. Afterwards, the upper and lower bounds of the correlation coefficient are defined. A theorem is given to determine these two bounds. It is stated that the correlation coefficient between two HFSs should also be hesitant, and thus, the upper and lower bounds can further help to identify the correlation coefficient between HFSs. The significant characteristic of the introduced correlation coefficient is that it lies in the interval $[-1, 1]$, which is in accordance with the classical correlation coefficient in statistics, whereas all the old correlation coefficients between HFSs in the literature are within the unit interval $[0, 1]$. The weighted correlation coefficient is also proposed to make it more applicable. In order to show the efficiency of the proposed correlation coefficients, they are implemented in medical diagnosis and cluster analysis. Some numerical examples are given in this chapter to illustrate the applicability and efficiency of the proposed correlation coefficient between HFSs.

Entropy is another important index for fuzzy information, which measures the degree of uncertainty of a fuzzy set. Usually, there are two aspects of uncertainty associated with a fuzzy set. One is related to fuzziness, which results from the lack

H. Liao and Z. Xu, *Hesitant Fuzzy Decision Making Methodologies and Applications*, Uncertainty and Operations Research,
DOI 10.1007/978-981-10-3265-3_2

of clear discrimination between the elements belonging or not belonging to a set. For classical fuzzy set, Zadeh (1965) first defined the entropy to measure the fuzziness of a fuzzy set and then many scholars developed different kinds of entropy formulas for fuzzy set (De Luca and Termini 1972; Kaufmann and Swanson 1975; Yager 1979; Parkash et al. 2008) and IFS (Burillo and Bustince 1996; Szmidt and Kacprzyk 2001; Wei et al. 2012). The other aspect of uncertainty associated with a fuzzy set is related to the lack of specificity. Specificity measures the amount of information contained in a fuzzy set. Yager (1992, 1998) put forward several specificity measures to quantify the degree that a fuzzy set contains just one element. Based on three t-norms and a negation, Garmendia et al. (2003) gave a general expression for specificity measures of a fuzzy set. Later, Yager (2008c) studied the formula of specificity measures in continuous domain. Pal et al. (2013) pointed out that there are two types of uncertainty for an IFS, i.e., the fuzzy-type uncertainty and the non-specificity type uncertainty. As to the entropy measure of HFS, Xu and Xia (2012) gave the axiomatic definition of entropy for HFEs and developed several entropy formulas to measure the degree of fuzziness of a HFE. Later, Farhadinia (2013) also proposed some entropy measures to quantify the degree of fuzziness of a HFS. In the second subsection of this chapter, we review the existing entropy measures for HFEs and demonstrate that the existing entropy measures for HFEs fail to effectively distinguish some apparently different HFEs in some cases. Then, we give a new axiomatic framework of entropy measures for HFEs by taking into account two facets of uncertainty associated with a HFE (i.e., fuzziness and non-specificity). We adopt a two-tuple entropy model to represent the two types of uncertainty associated with a HFE. Additionally, we discuss how to formulate each kind of uncertainty. Several examples are given to illustrate each method, and the comparisons with the existing entropy measures are also offered.

2.1 Novel Correlation Measures of Hesitant Fuzzy Sets

2.1.1 The Existing Correlation Measures of Hesitant Fuzzy Sets

As the correlation measure is one of the most important indices in measuring the relationship between two sets, it has been investigated in-depth within in the context of fuzzy sets and their extensions. As a representation, in the following, we just review the advances in correlation coefficient related to fuzzy sets and IFSs.

After discussing various properties which are attributed to "correlation" in statistics, Murthy et al. (1985) first introduced the correlation coefficient $\rho(\mu_1, \mu_2)$, similar to the correlation coefficient in statistics, between two fuzzy membership functions. It was proven that the correlation coefficient they defined satisfies many good properties, including $\rho(\mu_1, \mu_2) \in [-1, 1]$. In the case that the elements of fuzzy sets are ranked in terms of memberships, Chaudhuri and Bhattacharya (2001) proposed a rank correlation coefficient for fuzzy sets and then compared it with Murthy et al. (1985)'s correlation coefficient formula. Also adopting the concepts from conventional statistics, Chiang and Lin (1999) derived another formula of correlation coefficient in the domain of fuzzy sets. All these three kinds of correlation coefficients over fuzzy sets lie in the interval [−1, 1] and have similar meaning as that in conventional statistics. On the other hand, Yu (1993) introduced quite different concepts of correlation and correlation coefficient to measure the interrelation between fuzzy numbers. The value of correlation coefficient he introduced is within the interval [0, 1]. That is to say, the correlation coefficient he proposed can only represent the strength of relationship between fuzzy sets, but cannot manifest the positive or negative correlation.

It is stated that all the above achievements calculate the correlation coefficient between fuzzy sets as a crisp number. By using the sup-min convolution, Liu and Kao (2002) proposed a mathematical programming approach to calculate the correlation coefficient as a fuzzy number. After that, by applying the T_w-based extension principle, Hong (2006) gave an exact solution of a fuzzy correlation coefficient without relying on programming.

Regarding to IFSs, many different forms of correlation measures have also been investigated. Hung (2001) proposed the correlation coefficient for IFSs from statistics point of view by considering the membership degree and non-membership degree as two separate fuzzy sets. After that, Szmidt and Kacprzyk (2010) extended his formula by taking the hesitant degrees of IFSs into account. Mitchell (2004) also proposed an improved version of correlation coefficient, in which he interpreted the IFSs as the ensembles of ordinary membership functions. As these correlation coefficients are motivated from traditional statistics, the correlation coefficients of IFSs they developed are within the interval [−1, 1]. On the other side, motivated by the information energy of a fuzzy set, Gerstenkorn and Manko (1991) developed a quite different form of correlation coefficient for IFSs. Further, Hong and Hwang (1995) extended this type of correlation coefficient into possibility space in which the set $\{x_i\}$ is an infinite universe of discourse. Moreover, Hung and Wu (2002) improved the correlation coefficient and introduced the so-called centroid-method-based correlation coefficient for IFSs. As these correlation coefficients cannot guarantee that the correlation coefficient between any two IFSs equals to one if and only if these two IFSs are the same, Xu (2006b) proposed a new form of correlation coefficient for IFSs and circumvented this weakness. It should be stated that all the correlation coefficients proposed in Gerstenkorn and Manko (1991), Hong and Hwang (1995), Huang and Wu (2002), and Xu (2006b) lie in the unit interval [0, 1].

Some scholars also proposed distinct correlation measures within the context of HFSs. Xu and Xia (2011b) proposed several correlation coefficients from the point of HFEs. For two HFEs $h_A = \{\gamma_{A1}, \gamma_{A2}, \ldots, \gamma_{Al_A}\}$ and $h_B = \{\gamma_{B1}, \gamma_{B2}, \ldots, \gamma_{Bl_B}\}$, it is possible that the values in h_A and h_B are out of order. In addition, the number of values, l_A and l_B, in different HFEs may be different. Thus, to introduce the definition of correlation coefficient between two HFEs, Xu and Xia (2011b) firstly supposed that the values in different HFEs were arranged in ascending order;

meanwhile, they also assumed that the HFEs have the same length l. Based on these two assumptions, they proposed five different kinds of correlation coefficients for HFEs. Here we just set out one as a representation (for more others, readers can refer to Xu and Xia (2011b)):

$$\rho(h_A, h_B) = \frac{\sum_{k=1}^{l} \gamma_{A\sigma(k)}\gamma_{B\sigma(k)}}{\left(\sum_{k=1}^{l} \gamma_{A\sigma(k)}^2 \sum_{k=1}^{l} \gamma_{B\sigma(k)}^2\right)^{1/2}} \tag{2.1}$$

where $\sigma : (1, 2, \ldots, l) \to (1, 2, \ldots, l)$ is a permutation satisfying $\gamma_{\sigma(k)} \leq \gamma_{\sigma(k+1)}$, $k = 1, 2, \ldots, l-1$. Although Xu and Xia (2011b) stated that $|c(h_A, h_B)| \leq 1$, it is obvious that $c(h_A, h_B) \geq 0$ as $\gamma \in [0,\ 1]$, which means $c(h_A, h_B) \in [0,\ 1]$.

Chen et al. (2013a) proposed a formula to calculate the correlation coefficient between two HFSs. Let X be a reference set, $A = \{h_A(x_i)\}$ and $B = \{h_B(x_i)\}$ $(i = 1, 2, \ldots, n)$ be two HFSs. As the values in HFEs are out of order, and the number of values in different HFEs may be different, in order to introduce the correlation coefficient between two HFSs, the following assumptions are given in advance:

- The values in a HFE are arranged in ascending order.
- The lengths of different HFEs are assumed to have equal length.

The first assumption is easy to be satisfied. For the second one, sometimes the cardinality of two HFEs are different. In such case, as to Chen et al. (2013a)'s method, we need to make the lengths of the two HFEs be the same. There are many different regulations to extend the shorter HFE to the same length as the longer one. The most representative regulations are the pessimistic principle and the optimistic principle. For two HFEs h_A and h_B, let $l = \max\{l_{h_A}, l_{h_B}\}$ where l_{h_A} and l_{h_B} are the number of values in h_A and h_B, respectively. When $l_{h_A} \neq l_{h_B}$, one can extend the short HFE by adding some values in it until it has the same length with the other. In terms of the pessimistic principle, the short HFE is extended by adding the minimum value in it until it has the same length with the other HFE; while as to the optimistic principle, the maximum value of the short HFE should be added till the HFE has the same length as the longer one. In Chen et al. (2013a)'s definition, they used the former case and thus the correlation coefficient between two HFSs was defined as:

$$\rho_*(A, B) = \frac{\sum_{i=1}^{n} \left(\frac{1}{l_i} \sum_{k=1}^{l_i} \gamma_{A\sigma(k)}(x_i) \cdot \gamma_{B\sigma(k)}(x_i)\right)}{\left[\sum_{i=1}^{n} \left(\frac{1}{l_i} \sum_{k=1}^{l_i} \gamma_{A\sigma(k)}^2(x_i)\right)\right]^{1/2} \cdot \left[\sum_{i=1}^{n} \left(\frac{1}{l_i} \sum_{k=1}^{l_i} \gamma_{B\sigma(k)}^2(x_i)\right)\right]^{1/2}} \tag{2.2}$$

where $\gamma_{A\sigma(k)}(x_i)$ and $\gamma_{B\sigma(k)}(x_i)$ are the kth value in $h_A(x_i)$ and $h_B(x_i)$.

In summary, the correlation coefficients defined as Eqs. (2.1) and (2.2) have a few weaknesses:

(1) In Eq. (2.1), the HFEs were assumed to have equal length. This is not in accordance with real cases because it is impossible to make sure that all HFEs have equal length. As to Eq. (2.2), the pessimistic (or optimistic) principle was applied to fill the short HFE with some artificial values. It should be pointed out that filling some artificial values into a HFE would change its original information.

The following two examples show that the extensional regulations used in Xu and Xia (2011b) and Chen et al. (2013a) in the process of defining correlation coefficients between two HFEs or HFSs are not reasonable:

Example 2.1 (*Liao et al.* 2015b) For two HFEs $h_1 = \{0.1, 0.3\}$ and $h_2 = \{0.1, 0.3, 0.8\}$. In order to calculate the correlation coefficient between h_1 and h_2 by Eq. (2.1), according to Xu and Xia (2011b)'s assumptions, we should firstly extend h_1 to make it have equal length with h_2. Suppose that the pessimistic principle is applied to h_1, i.e., the minimum element in h_1 should be added to h_1. Then, h_1 is modified as $h_1' = \{0.1, 0.1, 0.3\}$. By Eqs. (1.7) and (1.8), we have $\bar{h}_1 = 0.2$, $\varphi_{h_1} = 0.1$, and $\bar{h}_1' = 0.1667$, $\varphi_{h_1'} = 0.0943$. It is obvious that the revised HFE h_1' is quite different from the original HFE h_1.

Example 2.2 (*Liao et al.* 2015b) Suppose that we are going to measure the correlation coefficient between two HFEs $h_1 = \{0.1, 0.8\}$ and $h_2 = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$. Then, according to the pessimistic principle, h_1 should be modified as $h_1' = \{0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.8\}$. Via Eqs. (1.7) and (1.8), we have $\bar{h}_1 = 0.45$, $\varphi_{h_1} = 0.35$, and $\bar{h}_1' = 0.1875$, $\varphi_{h_1'} = 0.2315$. It should be noted that the mean of the revised HFE h_1' is more than two times smaller than that of the original HFE h_1; meanwhile, the hesitant degree also changes apparently.

Examples 2.1 and 2.2 reveal that adding some artificial values in a HFE, no matter by pessimistic principle or by optimistic principle, would change the information of the original one. Thus, it is not very reasonable to measure the correlation coefficient between HFEs or HFSs by Eq. (2.1) or Eq. (2.2), and some new correlation coefficients need to be proposed for HFSs.

(2) The correlation coefficient defined in Xu and Xia (2011b) and Chen et al. (2013a) is always positive but this ignores the negative situation. In traditional random variable case, the correlation coefficient lies in $[-1, 1]$. For those correlation coefficients defined over fuzzy sets or IFSs, the correlation coefficients also lie in the interval $[-1, 1]$. Hence, it is not adequate to use the always positive variable to denote the correlation degree between two HFSs. The positive correlation coefficient can only demonstrate the strength of the relationship between HFSs, but cannot manifest the positive or negative correlation.
(3) It is not a best choice to use just one crisp number to represent the correlation degree between two HFEs or HFSs as the HFEs or HFSs per se are hesitant but

not precise. In other words, the correlation coefficient for HFSs should have certain degree of hesitance rather than just a fixed value.

2.1.2 Novel Correlation Measures of Hesitant Fuzzy Sets

This subsection introduces some novel correlation coefficients for HFSs. As to HFS, the following definitions are given:

Definition 2.1 (*Liao et al.* 2015b). For a reference set X, let $A = \{<x, h_A(x_i)> | x_i \in X\}$ be a HFS on X with $h_A(x_i) = \{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\}$, $i = 1, 2, \ldots, n$. The mean of the HFS A is defined as:

$$\bar{A} = E(A) = \frac{1}{n}\sum_{i=1}^{n} \bar{h}_A(x_i) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{l_{Ai}}\sum_{k=1}^{l_{Ai}} \gamma_{Aik}\right) \tag{2.3}$$

Definition 2.2 (*Liao et al.* 2015b). For a reference set X, let $A = \{<x, h_A(x_i)> | x_i \in X\}$ be a HFS on X with $h_A(x_i) = \{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\}$, $i = 1, 2, \ldots, n$. The variance of the HFS A is defined as:

$$Var(A) = \frac{1}{n}\sum_{i=1}^{n} (\bar{h}_A(x_i) - \bar{A}) \tag{2.4}$$

Definition 2.3 (*Liao et al.* 2015b). For a reference set X, let $A = \{h_A(x_i)\}$ and $B = \{h_B(x_i)\}$ be two HFSs on X, where $h_A(x_i) = \{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\}$, $h_B(x_i) = \{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\}$, $i = 1, 2, \ldots, n$. Then the correlation between HFSs A and B is defined as:

$$\begin{aligned} C(A,B) &= \frac{1}{n}\sum_{i=1}^{n} [\bar{h}_A(x_i) - \bar{A}] \cdot [\bar{h}_B(x_i) - \bar{B}] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[\bar{h}_A(x_i) - \frac{1}{n}\sum_{i=1}^{n}\bar{h}_A(x_i)\right] \cdot \left[\bar{h}_B(x_i) - \frac{1}{n}\sum_{i=1}^{n}\bar{h}_B(x_i)\right] \end{aligned} \tag{2.5}$$

where

$$\bar{h}_A(x_i) = \frac{1}{l_{Ai}}\sum_{k=1}^{l_{Ai}} \gamma_{Aik},\ \bar{h}_B(x_i) = \frac{1}{l_{Bi}}\sum_{k=1}^{l_{Bi}} \gamma_{Bik},\ i = 1, 2, \ldots, n \tag{2.6}$$

Note 2.1 In Definition 2.3, we do not need the HFSs A and B to have the same length, that is to say, $l_{Ai} \neq l_{Bi}$ is acceptable.

Note 2.2 The correlation above can be positive or negative. So the negative correlation can be modeled too.

Based on Definitions 2.2 and 2.3, it is easy to verify that the correlation coefficient between HFSs satisfies the following theorem:

Theorem 2.1 (Liao et al. 2015b). *For a HFS* $A = \{<x, h(x_i)> | x_i \in X\}$ *on* X *with* $h(x_i) = \{\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{il_i}\}$, $i = 1, 2, \ldots, n$, *the following equation holds:*

$$C(A, A) = Var(A) \tag{2.7}$$

Now we can define the correlation coefficient for HFSs:

Definition 2.4 (*Liao et al.* 2015b). For a reference set X, let $A = \{h_A(x_i)\}$ and $B = \{h_B(x_i)\}$ be two HFSs on X, where $h_A(x_i) = \{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}},\}$ and $h_B(x_i) = \{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\}$, $i = 1, 2, \ldots, n$. Then the correlation coefficient between the HFSs A and B is defined as:

$$\rho(A, B) = \frac{C(A, B)}{[C(A, A) \cdot C(B, B)]^{1/2}} \tag{2.8}$$

Theorem 2.2 reveals the fundamental properties of correlation coefficient between HFSs:

Theorem 2.2 (Liao et al. 2015b) *The correlation coefficient* $\rho(A, B)$ *between HFSs* A *and* B *satisfies the following properties:*

(1) $\rho(A, B) = \rho(B, A)$.
(2) $\rho(A, A) = 1$.
(3) $\rho(A, A^c) = -1$, *where* A^c *is defined as* $A^c = \{<x, h^c(x_i)> | x_i \in X\}$ *with* $h^c(x_i) = \{1 - \gamma_{i1}, 1 - \gamma_{i2}, \ldots, 1 - \gamma_{il_i}\}, i = 1, 2, \ldots, n$.
(4) $-1 \leq \rho(A, B) \leq 1$.

Proof The proofs of (1), (2) and (3) are obvious according to Definition 2.4. (4) According to Eq. (2.5), we have

$$|C(A, B)| = \left| \frac{1}{n} \sum_{i=1}^{n} [\bar{h}_A(x_i) - \bar{A}] \cdot [\bar{h}_B(x_i) - \bar{B}] \right| \leq \frac{1}{n} \sum_{i=1}^{n} \left| \bar{h}_A(x_i) - \bar{A} \right| \cdot \left| \bar{h}_B(x_i) - \bar{B} \right|$$

Using the Cauchy-Schwarz inequality:

$$(a_1 b_1 + a_2 b_2 + \cdots + a_n b_n)^2 \leq \left(a_1^2 + a_2^2 + \cdots + a_n^2\right) \cdot \left(b_1^2 + b_2^2 + \cdots + b_n^2\right)$$

where $a_i, b_i \in R, i = 1, 2, \ldots, N$, it follows that

$$
\begin{aligned}
|C(A,B)| &\le \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\bar{h}_A(x_i)-\bar{A}\right|^2}\cdot\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\bar{h}_B(x_i)-\bar{B}\right|^2} \\
&= [C(A,A)]^{1/2}\cdot[C(B,B)]^{1/2}
\end{aligned}
$$

Thus, we have

$$
|\rho(A,B)| = \frac{|C(A,B)|}{[C(A,A)]^{1/2}\cdot[C(B,B)]^{1/2}} \le 1
$$

Hence, $-1 \le \rho(A,B) \le 1$, which ends the proof.

In the following, we discuss how to measure the hesitant degree of the correlation coefficient $\rho(A,B)$ between two HFSs A and B.

We first rewrite the definition of correlation coefficient $\rho(A,B)$ between two HFSs A and B as:

$$
\rho(A,B) = \frac{\sum_{i=1}^{n}\left(\bar{h}_A(x_i)-\bar{A}\right)\cdot\left(\bar{h}_B(x_i)-\bar{B}\right)}{\sqrt{\sum_{i=1}^{n}\left(\bar{h}_A(x_i)-\bar{A}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\bar{h}_B(x_i)-\bar{B}\right)^2}} \tag{2.9}
$$

Then we define the upper and lower bounds of the correlation coefficient $\rho(A,B)$ between A and B as follows:

$$
\rho^U(A,B) = \max_{\substack{u_i \in h_A(x_i) \\ v_i \in h_B(x_i) \\ i=1,2,\ldots,n}} \frac{\sum_{i=1}^{n}(u_i-\bar{A})\cdot(v_i-\bar{B})}{\sqrt{\sum_{i=1}^{n}(u_i-\bar{A})^2}\cdot\sqrt{\sum_{i=1}^{n}(v_i-\bar{B})^2}} \tag{2.10}
$$

$$
\rho^L(A,B) = \min_{\substack{u_i \in h_A(x_i) \\ v_i \in h_B(x_i) \\ i=1,2,\ldots,n}} \frac{\sum_{i=1}^{n}(u_i-\bar{A})\cdot(v_i-\bar{B})}{\sqrt{\sum_{i=1}^{n}(u_i-\bar{A})^2}\cdot\sqrt{\sum_{i=1}^{n}(v_i-\bar{B})^2}} \tag{2.11}
$$

Theorem 2.3 (Liao et al. 2015b). *For two HFSs* $A = \{h_A(x_i)\}$ *and* $B = \{h_B(x_i)\}$ *on* X *with* $h_A(x_i) = \{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ai1_{Ai}}\}$, $h_B(x_i) = \{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bi1_{Bi}}\}$, $i = 1,2,\ldots,n$, *let* $h_A^U(x_i) = \max\{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\}$, $h_A^L(x_i) = \min\{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\}$, $h_B^U(x_i) = \max\{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\}$ *and* $h_B^L(x_i) = \min\{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\}$. *Then,*

(1) $\rho^L(A,B) \leq \rho(A,B) \leq \rho^U(A,B)$.

(2) $\rho^U(A,B) = \frac{\sum_{i=1}^n \left(h_A^U(x_i)-\bar{A}\right)\cdot\left(h_B^U(x_i)-\bar{B}\right)}{\sqrt{\sum_{i=1}^n \left(h_A^U(x_i)-\bar{A}\right)^2}\cdot\sqrt{\sum_{i=1}^n \left(h_B^U(x_i)-\bar{B}\right)^2}}$.

(3) $\rho^L(A,B) = \frac{\sum_{i=1}^n \left(h_A^L(x_i)-\bar{A}\right)\cdot\left(h_B^L(x_i)-\bar{B}\right)}{\sqrt{\sum_{i=1}^n \left(h_A^L(x_i)-\bar{A}\right)^2}\cdot\sqrt{\sum_{i=1}^n \left(h_B^L(x_i)-\bar{B}\right)^2}}$.

Note 2.3 Once we know $\rho^U(A,B)$ and $\rho^L(A,B)$, the difference

$$\Delta\rho(A,B) = \rho^U(A,B) - \rho^L(A,B) \tag{2.12}$$

can be used as an indicator how hesitant the correlation relationship is. The lager $\Delta\rho(A,B)$ is, the more hesitant the decision maker should be.

To prove the above theorem, a lemma is given first:

Lemma 2.1 (*Liao et al.* 2015b). Let x be any real number, $f(x) = \frac{x}{\sqrt{x^2+a}}$ with $a > 0$. Then, the function $f(x)$ is monotonically increasing.

Proof To prove the above, we only need to prove that $f'(x) > 0$. Since

$$f'(x) = \frac{\sqrt{x^2+a} - x\cdot\frac{2x}{2\sqrt{x^2+a}}}{x^2+a} = \frac{x^2+a-x^2}{(x^2+a)\cdot\sqrt{x^2+a}} = \frac{a}{(x^2+a)\cdot\sqrt{x^2+a}} > 0$$

then $f(x)$ is monotonically increasing, which competes the proof.

Based on the above lemma, we know for any real numbers x and y, if $x \geq y$, then

$$\frac{x}{\sqrt{x^2+a}} \geq \frac{y}{\sqrt{y^2+a}},\ a > 0 \tag{2.13}$$

In the following, we give the proof of Theorem 2.3.

Proof We only prove the case of $\rho^U(A,B)$. The case for $\rho^L(A,B)$ can be proven similarly.

Let $p_i = h_A^U(x_i) - \bar{A}$ and $q_i = u_i - \bar{A}$. Since $h_A^U(x_i) = \max\{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\}$ and $u_i \in h_A(x_i) = \{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\}$, then we get $h_A^U(x_i) \geq u_i$. Thus, $p_i \geq q_i$. Furthermore, we let $a_i = a_i(u_1, u_2, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n) \triangleq \sum_{\substack{j=1 \\ j\neq i}}^n \left(u_j - \bar{A}\right)^2$. Obviously, $a_i > 0$.

Analogously, for the HFS B, let $s_i = h_B^U(x_i) - \bar{B}$, $t_i = v_i - \bar{B}$, $b_i = b_i(v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \triangleq \sum_{\substack{j=1 \\ j\neq i}}^n \left(v_j - \bar{B}\right)^2$. We can also obtain $s_i \geq t_i$ as well as $b_i > 0$.

Based on the above transformation, the following equation holds:

$$\frac{\sum_{i=1}^{n}\left(h_A^U(x_i)-\bar{A}\right)\cdot\left(h_B^U(x_i)-\bar{B}\right)}{\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(u_j-\bar{A}\right)^2+\left(h_A^U(x_i)-\bar{A}\right)^2}\cdot\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(v_j-\bar{B}\right)^2+\left(h_B^U(x_i)-\bar{B}\right)^2}}$$
$$=\frac{\sum_{i=1}^{n}p_i\cdot s_i}{\sqrt{a_i+p_i^2}\cdot\sqrt{b_i+s_i^2}}$$

According to Eq. (2.13), it follows

$$\frac{\sum_{i=1}^{n}p_i\cdot s_i}{\sqrt{a_i+p_i^2}\cdot\sqrt{b_i+s_i^2}}=\sum_{i=1}^{n}\frac{p_i}{\sqrt{a_i+p_i^2}}\cdot\frac{s_i}{\sqrt{b_i+s_i^2}}$$
$$\geq\sum_{i=1}^{n}\frac{q_i}{\sqrt{a_i+q_i^2}}\cdot\frac{t_i}{\sqrt{b_i+t_i^2}}=\frac{\sum_{i=1}^{n}q_i\cdot t_i}{\sqrt{a_i+q_i^2}\cdot\sqrt{b_i+t_i^2}}$$
$$=\frac{\sum_{i=1}^{n}(u_i-\bar{A})\cdot(v_i-\bar{B})}{\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(u_j-\bar{A}\right)^2+(u_i-\bar{A})^2}\cdot\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(v_j-\bar{B}\right)^2+(v_i-\bar{B})^2}}=\frac{\sum_{i=1}^{n}(u_i-\bar{A})\cdot(v_i-\bar{B})}{\sqrt{\sum_{i=1}^{n}(u_i-\bar{A})^2}\cdot\sqrt{\sum_{i=1}^{n}(v_i-\bar{B})^2}}$$

Therefore, we can obtain

$$\frac{\sum_{i=1}^{n}\left(h_A^U(x_i)-\bar{A}\right)\cdot\left(h_B^U(x_i)-\bar{B}\right)}{\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(u_j-\bar{A}\right)^2+\left(h_A^U(x_i)-\bar{A}\right)^2}\cdot\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(v_j-\bar{B}\right)^2+\left(h_B^U(x_i)-\bar{B}\right)^2}}$$
$$\geq\frac{\sum_{i=1}^{n}(u_i-\bar{A})\cdot(v_i-\bar{B})}{\sqrt{\sum_{i=1}^{n}(u_i-\bar{A})^2}\cdot\sqrt{\sum_{i=1}^{n}(v_i-\bar{B})^2}},\ \text{for all}\ u_i\in\xi(x_i),\ v_i\in\xi(y_i),\ i=1,2,\cdots,n \tag{2.14}$$

In the left side of Eq. (2.14), we set $v_i=h_B^U(x_i)$. Then, it comes

$$\frac{\sum_{i=1}^{n}\left(h_A^U(x_i)-\bar{A}\right)\cdot\left(h_B^U(x_i)-\bar{B}\right)}{\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(u_j-\bar{A}\right)^2+\left(h_A^U(x_i)-\bar{A}\right)^2}\cdot\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(v_j-\bar{B}\right)^2+\left(h_B^U(x_i)-\bar{B}\right)^2}}$$
$$=\frac{\sum_{i=1}^{n}\left(h_A^U(x_i)-\bar{A}\right)\cdot\left(h_B^U(x_i)-\bar{B}\right)}{\sqrt{\sum_{i=1}^{n}\left(h_A^U(x_i)-\bar{A}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(h_B^U(x_i)-\bar{B}\right)^2}}\geq\frac{\sum_{i=1}^{n}(u_i-\bar{A})\cdot(v_i-\bar{B})}{\sqrt{\sum_{i=1}^{n}(u_i-\bar{A})^2}\cdot\sqrt{\sum_{i=1}^{n}(v_i-\bar{B})^2}},$$

$$\text{for all}\ u_i\in h_A(x_i),\ v_i\in h_B(x_i),\ i=1,2,\ldots,n \tag{2.15}$$

Combining Eqs. (2.9) and (2.15), we can obtain

$$\rho^U(A,B) = \frac{\sum_{i=1}^{n}\left(h_A^U(x_i) - \bar{A}\right)\cdot\left(h_B^U(x_i) - \bar{B}\right)}{\sqrt{\sum_{i=1}^{n}\left(h_A^U(x_i) - \bar{A}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(h_B^U(x_i) - \bar{B}\right)^2}}$$

which is to say, (2) in Theorem 2.3 holds.

Additionally, in the right side of the inequality (2.15), let $v_i = \bar{h}_B(x_i)$, then it follows

$$\frac{\sum_{i=1}^{n}\left(h_A^U(x_i) - \bar{A}\right)\cdot\left(h_B^U(x_i) - \bar{B}\right)}{\sqrt{\sum_{i=1}^{n}\left(h_A^U(x_i) - \bar{A}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(h_B^U(x_i) - \bar{B}\right)^2}} \geq \frac{\sum_{i=1}^{n}\left(\bar{h}_A(x_i) - \bar{A}\right)\cdot\left(\bar{h}_B(x_i) - \bar{B}\right)}{\sqrt{\sum_{i=1}^{n}\left(\bar{h}_A(x_i) - \bar{A}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(\bar{h}_B(x_i) - \bar{B}\right)^2}}$$

i.e.,

$$\rho^U(A,B) \geq \rho(A,B)$$

This completes the proof.

Definition 2.5 (*Liao et al.* 2015b). For a reference set X, let $A = \{h_A(x_i)\}$ and $B = \{h_B(x_i)\}$ be two HFSs on X, where $h_A(x_i) = \left\{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\right\}$, and $h_B(x_i) = \left\{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\right\}$, $i = 1, 2, \ldots, n$. Given the correlation coefficient $\rho(A,B)$ between A and B being defined as Eq. (2.9), then the hesitant degree of $\rho(A,B)$ is measured in terms of

$$\varphi_{(A,B)} = \rho^U(A,B) - \rho^L(A,B) \tag{2.16}$$

where $\rho^U(A,B)$ and $\rho^L(A,B)$ are the upper and lower bounds of the correlation coefficient $\rho(A,B)$.

It is noted that the value of correlation coefficient $\rho(A,B)$ between the HFSs A and B defined as Eq. (2.8) is also a crisp value. However, as both A and B are HFSs, it is not adequate to use just a crisp value to represent their relationship. The correlation coefficient defined as Eq. (2.8) can only be taken as the expected (or averaging) correlation coefficient between the HFSs A and B. In order to describe the correlation coefficient between HFSs more objectively, we can also use the upper bound $\rho^U(A,B)$, the lower bound $\rho^L(A,B)$, or the hesitant degree $\varphi_{(A,B)}$ to better identify the correlation coefficient between two HFSs A and B.

Consider that in some cases, the objects $x_i \in X\ (i = 1, 2, \ldots, n)$ may be assigned different weights. Liao et al. (2015b) proposed the weighted form of the correlation coefficient for HFSs.

Definition 2.6 (*Liao et al.* 2015b). Let $w = (w_1, w_2, \ldots, w_n)$ be the weight vector of $x_i\ (i = 1, 2, \ldots, n)$ with $w_i \in [0,\ 1]$, $(i = 1, 2, \ldots, n)$ and $\sum_{i=1}^{n} w_i = 1$. For two HFSs $A = \{h_A(x_i)\}$ and $B = \{h_B(x_i)\}$ with $h_A(x_i) = \left\{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\right\}$, and

$h_B(x_i) = \{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\}$, $i = 1, 2, \ldots, n$, the following definition can be developed:

(1) The weighted mean of the HFS A is defined as:

$$\bar{A}_w = \frac{1}{n}\sum_{i=1}^{n} w_i \bar{h}_A(x_i) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{w_i}{l_{Ai}}\sum_{k=1}^{l_{Ai}}\gamma_{Aik}\right) \tag{2.17}$$

(2) The weighted variance of the HFS A is defined as:

$$Var_w(A) = \frac{1}{n}\sum_{i=1}^{n}\left(w_i \bar{h}_A(x_i) - \bar{A}_w\right) \tag{2.18}$$

(3) The weighted correlation between the HFSs A and B is defined as:

$$\begin{aligned} C_w(A,B) &= \frac{1}{n}\sum_{i=1}^{n}\left[w_i \bar{h}_A(x_i) - \bar{A}_w\right]\cdot\left[w_i \bar{h}_B(x_i) - \bar{B}_w\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[w_i \bar{h}_A(x_i) - \frac{1}{n}\sum_{i=1}^{n} w_i \bar{h}_A(x_i)\right]\cdot\left[w_i \bar{h}_B(x_i) - \frac{1}{n}\sum_{i=1}^{n} w_i \bar{h}_B(x_i)\right] \end{aligned} \tag{2.19}$$

where
$\bar{h}_A(x_i) = \frac{1}{l_{Ai}}\sum_{k=1}^{l_{Ai}}\gamma_{Aik}$, $\bar{h}_B(x_i) = \frac{1}{l_{Bi}}\sum_{k=1}^{l_{Bi}}\gamma_{Bik}$, $i = 1, 2, \ldots, n$

(4) The weighted correlation coefficient between the HFSs A and B is defined as:

$$\begin{aligned} \rho_w(A,B) &= \frac{C_w(A,B)}{\left[C_w(A,A)\cdot C_w(B,B)\right]^{1/2}} \\ &= \frac{\sum_{i=1}^{n}\left(w_i \bar{h}_A(x_i) - \bar{A}_w\right)\cdot\left(w_i \bar{h}_B(x_i) - \bar{B}_w\right)}{\sqrt{\sum_{i=1}^{n}\left(w_i \bar{h}_A(x_i) - \bar{A}_w\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(w_i \bar{h}_B(x_i) - \bar{B}_w\right)^2}} \end{aligned} \tag{2.20}$$

The weighted correlation coefficient $\rho_w(A,B)$ between the HFSs A and B also satisfies the following properties:

(a) $\rho_w(A,B) = \rho_w(B,A)$;
(b) $\rho_w(A,A) = 1$;
(c) $\rho_w(A,A^c) = -1$, where A^c is defined as $A^c = \{<x, h^c(x_i)> | x_i \in X\}$ with $h^c(x_i) = \{1-\gamma_{i1}, 1-\gamma_{i2}, \ldots, 1-\gamma_{il_i}\}$, $i = 1, 2, \ldots, n$;
(d) $-1 \le \rho_w(A,B) \le 1$.

(5) The upper and lower bounds of the weighted correlation coefficient $\rho_w(A,B)$ are defined as:

$$\rho_w^U(A,B) = \max_{\substack{u_i \in h_A(x_i) \\ v_i \in h_B(x_i) \\ i=1,2,\ldots,n}} \frac{\sum_{i=1}^n \left(w_i u_i - \bar{A}_w\right)\cdot\left(w_i v_i - \bar{B}_w\right)}{\sqrt{\sum_{i=1}^n \left(w_i u_i - \bar{A}_w\right)^2}\cdot\sqrt{\sum_{i=1}^n \left(w_i v_i - \bar{B}_w\right)^2}} \tag{2.21}$$

$$\rho_w^L(A,B) = \min_{\substack{u_i \in h_A(x_i) \\ v_i \in h_B(x_i) \\ i=1,2,\ldots,n}} \frac{\sum_{i=1}^n \left(w_i u_i - \bar{A}_w\right)\cdot\left(w_i v_i - \bar{B}_w\right)}{\sqrt{\sum_{i=1}^n \left(w_i u_i - \bar{A}_w\right)^2}\cdot\sqrt{\sum_{i=1}^n \left(w_i v_i - \bar{B}_w\right)^2}} \tag{2.22}$$

The following properties hold as well:

(a) $$\rho_w^L(A,B) \le \rho_w(A,B) \le \rho_w^U(A,B).$$

(b) $$\rho_w^U(A,B) = \frac{\sum_{i=1}^n \left(w_i h_A^U(x_i) - \bar{A}_w\right)\cdot\left(w_i h_B^U(x_i) - \bar{B}_w\right)}{\sqrt{\sum_{i=1}^n \left(w_i h_A^U(x_i) - \bar{A}_w\right)^2}\cdot\sqrt{\sum_{i=1}^n \left(w_i h_B^U(x_i) - \bar{B}_w\right)^2}}.$$

(c) $\rho_w^L(A,B) = \frac{\sum_{i=1}^n \left(w_i h_A^L(x_i) - \bar{A}_w\right)\cdot\left(w_i h_B^L(x_i) - \bar{B}_w\right)}{\sqrt{\sum_{i=1}^n \left(w_i h_A^L(x_i) - \bar{A}_w\right)^2}\cdot\sqrt{\sum_{i=1}^n \left(w_i h_B^L(x_i) - \bar{B}_w\right)^2}}$, where

$$h_A^U(x_i) = \max\left\{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\right\},\ h_A^L(x_i) = \min\left\{\gamma_{Ai1}, \gamma_{Ai2}, \ldots, \gamma_{Ail_{Ai}}\right\}$$
$$h_B^U(x_i) = \max\left\{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\right\},\ h_B^L(x_i) = \min\left\{\gamma_{Bi1}, \gamma_{Bi2}, \ldots, \gamma_{Bil_{Bi}}\right\}$$

(6) The hesitant degree of $\rho_w(A,B)$ is measured in terms of

$$\varphi_{(A,B)_w} = \rho_w^U(A,B) - \rho_w^L(A,B) \tag{2.23}$$

2.1.3 *Applications of the Correlation Measures of Hesitant Fuzzy Sets*

(1) The application of the correlation coefficients in medical diagnosis

The correlation coefficient can be implemented into many practical applications. The first case given below is related to medical diagnosis.

Example 2.3 (*Liao et al.* 2015b). Suppose that a doctor wants to make a proper diagnosis D = {Viral fever, Malaria, Typhoid, Stomach problem, Chest problem} for a set of patients P = {Al, Bob, Joe, Ted} with the values of symptoms

Table 2.1 Symptom characteristics for the considered diagnoses in terms of HFSs

	Temperature	Headache	Cough	Stomach pain	Stomach pain
Viral fever	$\{0.6, 0.4, 0.3\}$	$\{0.7, 0.5, 0.3, 0.2\}$	$\{0.5, 0.3\}$	$\{0.5, 0.4, 0.3, 0.2, 0.1\}$	$\{0.5, 0.4, 0.2, 0.1\}$
Malaria	$\{0.9, 0.8, 0.7\}$	$\{0.5, 0.3, 0.2, 0.1\}$	$\{0.2, 0.1\}$	$\{0.6, 0.5, 0.3, 0.2, 0.1\}$	$\{0.4, 0.3, 0.2, 0.1\}$
Typhoid	$\{0.6, 0.3, 0.1\}$	$\{0.9, 0.8, 0.7, 0.6\}$	$\{0.5, 0.3\}$	$\{0.5, 0.4, 0.3, 0.2, 0.1\}$	$\{0.6, 0.4, 0.3, 0.1\}$
Stomach problem	$\{0.5, 0.4, 0.2\}$	$\{0.4, 0.3, 0.2, 0.1\}$	$\{0.4, 0.3\}$	$\{0.9, 0.8, 0.7, 0.6, 0.5\}$	$\{0.5, 0.4, 0.2, 0.1\}$
Chest problem	$\{0.3, 0.2, 0.1\}$	$\{0.5, 0.3, 0.2, 0.1\}$	$\{0.3, 0.2\}$	$\{0.7, 0.6, 0.5, 0.3, 0.2\}$	$\{0.9, 0.8, 0.7, 0.6\}$

Table 2.2 Symptom characteristics for the considered patients in terms of HFSs

	Temperature	Headache	Cough	Stomach pain	Chester pain
Al	$\{0.9, 0.7, 0.5\}$	$\{0.4, 0.3, 0.2, 0.1\}$	$\{0.4, 0.3\}$	$\{0.6, 0.5, 0.4, 0.2, 0.1\}$	$\{0.4, 0.3, 0.2, 0.1\}$
Bob	$\{0.5, 0.4, 0.2\}$	$\{0.5, 0.4, 0.3, 0.1\}$	$\{0.2, 0.1\}$	$\{0.9, 0.8, 0.6, 0.5, 0.4\}$	$\{0.5, 0.4, 0.3, 0.2\}$
Joe	$\{0.9, 0.7, 0.6\}$	$\{0.7, 0.4, 0.3, 0.1\}$	$\{0.3, 0.2\}$	$\{0.6, 0.4, 0.3, 0.2, 0.1\}$	$\{0.6, 0.3, 0.2, 0.1\}$
Ted	$\{0.8, 0.7, 0.5\}$	$\{0.6, 0.5, 0.4, 0.2\}$	$\{0.4, 0.3\}$	$\{0.6, 0.4, 0.3, 0.2, 0.1\}$	$\{0.5, 0.4, 0.2, 0.1\}$

V = {temperature, headache, cough, stomach pain, chest pain}. As in many cases such as in traditional Chinese medical diagnosis or in emergency case that crisp measuring instruments cannot be obtained, it is impossible to get the crisp values of the symptoms but only vague information, which is described in terms of HFEs. Before starting the diagnosis, a medical knowledge-based data set involving symptom characteristic of the considered diagnoses is necessary to be constructed (see Table 2.1). The symptoms of the patients are given in Table 2.2.

To derive a diagnosis for each patient, we can calculate the correlation coefficient between the symptom characteristic of each diagnose and that of each patient. Using the correlation coefficient formula shown as Eq. (2.9), the correlation coefficient values are obtained, shown in Table 2.3 and Fig. 2.1. From Table 2.3 and Fig. 2.1, it is clear to see that Al, Joe and Ted suffer from Malaria, but Bob suffers from Stomach problem.

Meanwhile, Xu and Xia (2011b) utilized the correlation formula Eq. (2.1) to calculate the correlation coefficients and yielded their results, illustrated in Table 2.4 and Fig. 2.2. Table 2.4 and Fig. 2.2 imply that Al and Ted suffer from viral fever; Bob suffer from stomach problem; Joe suffer from malaria.

Table 2.3 Correlation coefficient values for each considered patient to the set of possible diagnoses by using our approach

	Viral fever	Malaria	Typhoid	Stomach problem	Chest problem
Al	0.4597	**0.9187**	−0.4288	0.1323	−0.5372
Bob	−0.5715	0.2546	−0.3166	**0.8074**	0.3042
Joe	0.5395	**0.9803**	−0.1217	−0.1017	−0.4636
Ted	0.7330	**0.9082**	0.0210	−0.2230	−0.6506

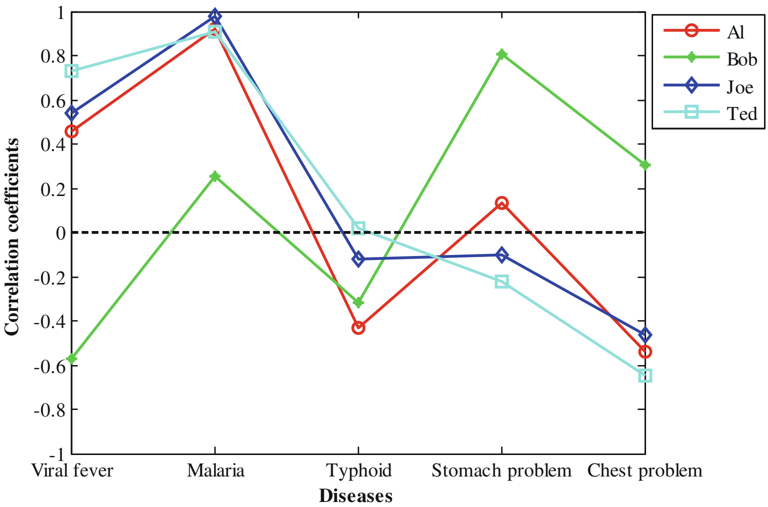

Fig. 2.1 Correlation coefficient values by using our approach

Table 2.4 Correlation coefficient values for each considered patient to the set of possible diagnoses by using Xu and Xia (2011b)'s approach

	Viral fever	Malaria	Typhoid	Stomach problem	Chest Problem
Al	**0.9969**	0.9929	0.9800	0.9902	0.9878
Bob	0.9900	0.9862	0.9792	**0.9921**	0.9909
Joe	0.9927	**0.9929**	0.9677	0.9817	0.9750
Ted	**0.9942**	0.9899	0.9787	0.9879	0.9772

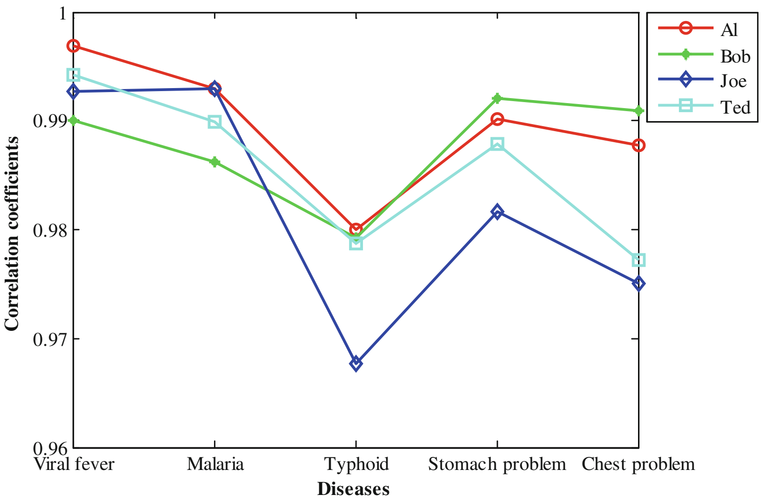

Fig. 2.2 Correlation coefficient values by using Xu and Xia (2011b)'s approach

Comparing the results in Table 2.3 with those in Table 2.4, some interesting findings can be derived. Firstly, we can find that all the values in Table 2.4 are positive values within the unit interval [0, 1], but in Table 2.3, there are some negative values. It is quite strange that all the HFEs are positively correlated even though all the values over different symptom characteristics are quite different. This is the first weakness of Xu and Xia (2011b)'s method. For example, let us look into the symptom characteristics of Typhoid and those of Al. It is obvious that Al's symptoms are negative correlated to those of Typhoid. However, according to Eq. (2.1), the correlation between Typhoid and Al is 0.9800, which implies that it is highly probable that Al suffers from Typhoid. This is definitely wrong.

In addition, comparing Table 2.3 (or Fig. 2.1) with Table 2.4 (or Fig. 2.2), we can find that all the correlation coefficients shown in Table 2.4 are quite close and vary from 0.9677 to 0.9969. These similar values cannot clearly distinguish the different between different diagnoses. Actually, if we draw a new figure (see Fig. 2.3) according to Xu and Xia (2011b)'s results but restrict the correlation coefficient values vary within the same domain as in Fig. 2.1, then it is very hard or even impossible for us to distinguish the diagnoses. In other words, the results derived from Table 2.4 are not very convincing (or at least not applicable) especially when the number of objects is a little large. However, Table 2.3 presents a striking contrast to Table 2.4 as all the values in it lies between −0.4288 and 0.9803, which shows the differences among the diagnoses significantly. All these above points imply that the correlation coefficient proposed in this chapter is much more convincing in medical diagnosis.

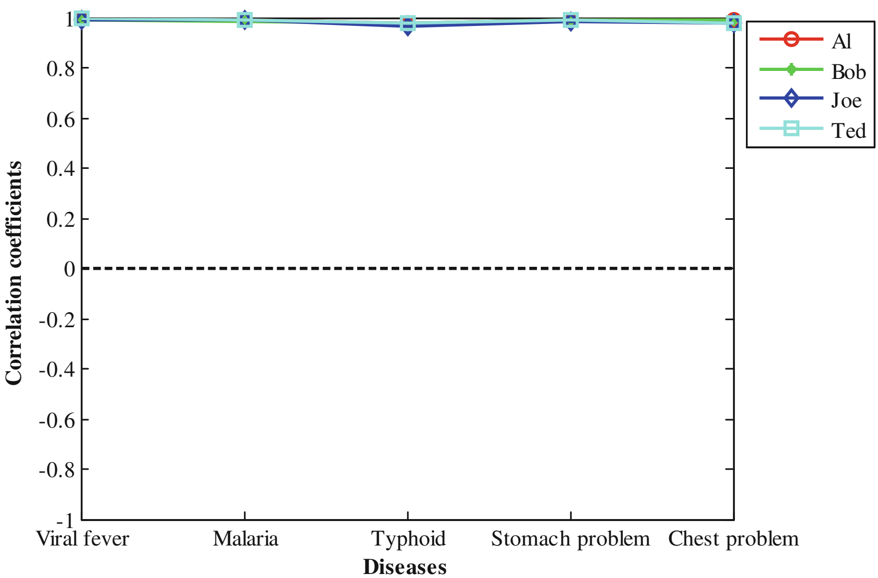

Fig. 2.3 Correlation coefficient values by using Xu and Xia (2011b)'s approach

It is stated that in this example, we just use Eq. (2.9) as a representation to describe the correlation coefficient between HFSs and illustrate its advantages over the existing correlation coefficients for HFSs. In fact, we can also use the upper bound $\rho^U(A,B)$, the lower bound $\rho^L(A,B)$, or the hesitant degree $\varphi_{(A,B)}$ to better identify the correlation coefficients in the above example.

(2) The application of the correlation coefficients in cluster analysis

To better understand the strength of the novel correlation coefficients, in the following, we show the applicability of the correlation coefficient between HFSs in the process of clustering. Cluster analysis, or clustering, is defined as the unsupervised process of group a set of data objects in such a way that objects in the same group (called a cluster) are somehow more similar to each other than to those in other groups (clusters) (Jain et al. 1999). It can be applied either as an exploratory tool (to discover previously unknown pattern in data), or as an input to a decision making process (Friedman et al. 2007). There are many algorithms for clustering, which differ significantly in their notion of what constitutes a cluster and how to efficiently find them. Within the context of hesitant fuzzy information, Chen et al. (2013a) proposed an algorithm to cluster hesitant fuzzy data into different clusters. In that algorithm, the correlation coefficient defined as Eq. (2.2) is used to measure the relationship between different objects. In the following, we do not intend to propose new clustering algorithm but use that algorithm to illustrate the efficiency of our proposed correlation coefficient. The algorithm proposed by Chen et al. (2013a) is described below:

Algorithm 2.1

Step 1. Let $\{A_1, A_2, \ldots, A_m\}$ be a set of HFSs on X. We construct a correlation matrix $C = (\rho_{ij})_{m\times m}$ where $\rho_{ij} = \rho(A_i, A_j)$ and can be calculated via Eq. (2.1) or Eq. (2.9) or Eq. (2.20).

Step 2. Check whether the correlation matrix satisfies $C^2 \subseteq C$, where $C^2 = C \circ C = \left(\rho'_{ij}\right)_{m\times m}$, and $\rho'_{ij} = \max_k\{\min\{\rho_{ik}, \rho_{kj}\}\}$, $i,j = 1,2,\ldots,m$. If it does not hold, then we construct the equivalent correlation matrix $C^{2^k}: C \to C^2 \to C^4 \to \cdots \to C^{2^k} \to \cdots$ until $C^{2^k} = C^{2^{(k+1)}}$.

Step 3. For a given confidence level $\lambda \in [0, 1]$, we construct a λ-cutting matrix $C_\lambda = \left(\rho^\lambda_{ij}\right)_{m\times m}$ where ρ^λ_{ij} is defined as:

$$\rho^\lambda_{ij} = \begin{cases} 0 & if\ \rho_{ij} < \lambda \\ 1 & if\ \rho_{ij} \geq \lambda \end{cases} \quad i,j = 1,2,\ldots,m \tag{2.24}$$

Step 4. Classify the HFEs by the principle: if all elements of the ith line in C_λ are the same as the corresponding elements of the jth line, then the HFEs A_i and A_j are supposed as the same type.

Step 5. End.

An application example concerning the assessment of business failure risk is utilized to validate the above algorithm and our proposed correlation coefficient for HFSs. In this example, the weighted correlation coefficient defined as Eq. (2.20) is used to measure the correlation coefficient between HFSs:

Example 2.4 (*Liao et al.* 2015b). The assessment of business failure risk, i.e., the assessment of firm performance and the prediction of failure events has drawn the attention of many researchers in recent years (Chen et al. 2013a). Suppose that there are 10 firms $A_i (i = 1, 2, \ldots, 10)$ to be evaluated by several risk evaluation organizations from different aspects. To get fair assessments for these firms, the risk evaluation organizations established five criteria: ζ_1 : managers work experience, ζ_2 : profitability, ζ_3 : operating capacity, ζ_4 : debt-paying ability, and ζ_5 : market competition, whose weighting vector is set as $w = (0.15, 0.3, 0.2, 0.25, 0.1)$. As the risk evaluation organizations have different backgrounds and knowledge, it is possible that they may get different evaluation values from their perspectives. To better reflect the opinions established by different organizations, the evaluation values given by them are represented by HFEs and displayed in Table 2.5.

In the following, we use Algorithm 2.1 and the weighted correlation coefficient to cluster the firms.

Step 1. By Eq. (2.20), we can calculate the weighted correlation coefficients between each pair of the alternatives $\rho(A_i, A_j)$, $i, j = 1, 2, \ldots, 10$:

$$C_w = \begin{pmatrix} 1.0000 & -0.8347 & -0.6840 & -0.0619 & -0.7198 & 0.8272 & 0.4225 & -0.6728 & -0.2983 & -0.3817 \\ -0.8347 & 1.0000 & 0.9659 & 0.5143 & 0.6062 & -0.8432 & 0.0948 & 0.5874 & 0.7511 & -0.1724 \\ -0.6840 & 0.9659 & 1.0000 & 0.6586 & 0.5097 & -0.7364 & 0.3472 & 0.5766 & 0.8665 & -0.3965 \\ -0.0619 & 0.5143 & 0.6586 & 1.0000 & -0.3041 & -0.0295 & 0.7073 & -0.0649 & 0.9365 & -0.8463 \\ -0.7198 & 0.6062 & 0.5097 & -0.3041 & 1.0000 & -0.8852 & -0.2776 & 0.8068 & 0.0388 & 0.3949 \\ 0.8272 & -0.8432 & -0.7364 & -0.0295 & -0.8852 & 1.0000 & 0.2176 & -0.6119 & -0.3648 & -0.1863 \\ 0.4225 & 0.0948 & 0.3472 & 0.7073 & -0.2776 & 0.2176 & 1.0000 & -0.0034 & 0.6454 & -0.9438 \\ -0.6728 & 0.5874 & 0.5766 & -0.0649 & 0.8068 & -0.6119 & -0.0034 & 1.0000 & 0.1741 & 0.2108 \\ -0.2983 & 0.7511 & 0.8665 & 0.9365 & 0.0388 & -0.3648 & 0.6454 & 0.1741 & 1.0000 & -0.7634 \\ -0.3817 & -0.1724 & -0.3965 & -0.8463 & 0.3949 & -0.1863 & -0.9438 & 0.2108 & -0.7634 & 1.0000 \end{pmatrix}$$

Table 2.5 The evaluation information for the 5 criteria of 10 firms

	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5
A_1	{0.3,0.4,0.5}	{0.4,0.5}	{0.8}	{0.5}	{0.2,0.3}
A_2	{0.4,0.6}	{0.6,0.8}	{0.2,0.3}	{0.3,0.4}	{0.6,0.7,0.9}
A_3	{0.5,0.7}	{0.9}	{0.3,0.4}	{0.3}	{0.8,0.9}
A_4	{0.3,0.4,0.5}	{0.8,0.9}	{0.7,0.9}	{0.1,0.2}	{0.9,1.0}
A_5	{0.8,1.0}	{0.8,1.0}	{0.4,0.6}	{0.8}	{0.7,0.8}
A_6	{0.4,0.5,0.6}	{0.2,0.3}	{0.9,1.0}	{0.5}	{0.3,0.4,0.5}
A_7	{0.6}	{0.7,0.9}	{0.8}	{0.3,0.4}	{0.4,0.7}
A_8	{0.9,1.0}	{0.7,0.8}	{0.4,0.5}	{0.5,0.6}	{0.7}
A_9	{0.4,0.6}	{1.0}	{0.6,0.7}	{0.2,0.3}	{0.9,1.0}
A_{10}	{0.9}	{0.6,0.7}	{0.5,0.8}	{1.0}	{0.7,0.8,0.9}

Step 2. The equivalent correlation matrix is constructed as follows:

$$C_w^2 = \begin{pmatrix} 1.0000 & 0.0948 & 0.3472 & 0.4225 & -0.2776 & 0.8272 & 0.4225 & -0.0034 & 0.4225 & -0.1863 \\ 0.0948 & 1.0000 & 0.9659 & 0.7511 & 0.6062 & 0.0948 & 0.6454 & 0.6062 & 0.8665 & 0.3949 \\ 0.3472 & 0.9659 & 1.0000 & 0.8665 & 0.6062 & 0.2176 & 0.6586 & 0.5874 & 0.8665 & 0.3949 \\ 0.4225 & 0.7511 & 0.8665 & 1.0000 & 0.5143 & 0.2176 & 0.7073 & 0.5766 & 0.9365 & -0.0649 \\ -0.2776 & 0.6062 & 0.6062 & 0.5143 & 1.0000 & -0.1863 & 0.3472 & 0.8068 & 0.6062 & 0.3949 \\ 0.8272 & 0.0948 & 0.2176 & 0.2176 & -0.1863 & 1.0000 & 0.4225 & -0.0034 & 0.2176 & -0.1863 \\ 0.4225 & 0.6454 & 0.6586 & 0.7073 & 0.3472 & 0.4225 & 1.0000 & 0.3472 & 0.7073 & -0.0034 \\ -0.0034 & 0.6062 & 0.5874 & 0.5766 & 0.8068 & -0.0034 & 0.3472 & 1.0000 & 0.5874 & 0.3949 \\ 0.4225 & 0.8665 & 0.8665 & 0.9365 & 0.6062 & 0.2176 & 0.7073 & 0.5874 & 1.0000 & 0.1741 \\ -0.1863 & 0.3949 & 0.3949 & -0.0649 & 0.3949 & -0.1863 & -0.0034 & 0.3949 & 0.1741 & 1.0000 \end{pmatrix}$$

$$C_w^4 = \begin{pmatrix} 1.0000 & 0.4225 & 0.4225 & 0.4225 & 0.4225 & 0.8272 & 0.4225 & 0.4225 & 0.4225 & 0.3472 \\ 0.4225 & 1.0000 & 0.9659 & 0.8665 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.8665 & 0.3949 \\ 0.4225 & 0.9659 & 1.0000 & 0.8665 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.8665 & 0.3949 \\ 0.4225 & 0.8665 & 0.8665 & 1.0000 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.9365 & 0.3949 \\ 0.4225 & 0.6062 & 0.6062 & 0.6062 & 1.0000 & 0.3472 & 0.6062 & 0.8068 & 0.6062 & 0.3949 \\ 0.8272 & 0.4225 & 0.4225 & 0.4225 & 0.3472 & 1.0000 & 0.4225 & 0.3472 & 0.4225 & 0.2176 \\ 0.4225 & 0.7073 & 0.7073 & 0.7073 & 0.6062 & 0.4225 & 1.0000 & 0.6062 & 0.7073 & 0.3949 \\ 0.4225 & 0.6062 & 0.6062 & 0.6062 & 0.8068 & 0.3472 & 0.6062 & 1.0000 & 0.6062 & 0.3949 \\ 0.4225 & 0.8665 & 0.8665 & 0.9365 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 1.0000 & 0.3949 \\ 0.3472 & 0.3949 & 0.3949 & -0.0649 & 0.3949 & 0.2176 & 0.3949 & 0.3949 & 0.3949 & 1.0000 \end{pmatrix}$$

$$C_w^8 = \begin{pmatrix} 1.0000 & 0.4225 & 0.4225 & 0.4225 & 0.4225 & 0.8272 & 0.4225 & 0.4225 & 0.4225 & 0.3949 \\ 0.4225 & 1.0000 & 0.9659 & 0.8665 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.8665 & 0.3949 \\ 0.4225 & 0.9659 & 1.0000 & 0.8665 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.8665 & 0.3949 \\ 0.4225 & 0.8665 & 0.8665 & 1.0000 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.9365 & 0.3949 \\ 0.4225 & 0.6062 & 0.6062 & 0.6062 & 1.0000 & 0.4225 & 0.6062 & 0.8068 & 0.6062 & 0.3949 \\ 0.8272 & 0.4225 & 0.4225 & 0.4225 & 0.4225 & 1.0000 & 0.4225 & 0.4225 & 0.4225 & 0.3949 \\ 0.4225 & 0.7073 & 0.7073 & 0.7073 & 0.6062 & 0.4225 & 1.0000 & 0.6062 & 0.7073 & 0.3949 \\ 0.4225 & 0.6062 & 0.6062 & 0.6062 & 0.8068 & 0.4225 & 0.6062 & 1.0000 & 0.6062 & 0.3949 \\ 0.4225 & 0.8665 & 0.8665 & 0.9365 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 1.0000 & 0.3949 \\ 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 1.0000 \end{pmatrix}$$

$$C_w^{16} = \begin{pmatrix} 1.0000 & 0.4225 & 0.4225 & 0.4225 & 0.4225 & 0.8272 & 0.4225 & 0.4225 & 0.4225 & 0.3949 \\ 0.4225 & 1.0000 & 0.9659 & 0.8665 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.8665 & 0.3949 \\ 0.4225 & 0.9659 & 1.0000 & 0.8665 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.8665 & 0.3949 \\ 0.4225 & 0.8665 & 0.8665 & 1.0000 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 0.9365 & 0.3949 \\ 0.4225 & 0.6062 & 0.6062 & 0.6062 & 1.0000 & 0.4225 & 0.6062 & 0.8068 & 0.6062 & 0.3949 \\ 0.8272 & 0.4225 & 0.4225 & 0.4225 & 0.4225 & 1.0000 & 0.4225 & 0.4225 & 0.4225 & 0.3949 \\ 0.4225 & 0.7073 & 0.7073 & 0.7073 & 0.6062 & 0.4225 & 1.0000 & 0.6062 & 0.7073 & 0.3949 \\ 0.4225 & 0.6062 & 0.6062 & 0.6062 & 0.8068 & 0.4225 & 0.6062 & 1.0000 & 0.6062 & 0.3949 \\ 0.4225 & 0.8665 & 0.8665 & 0.9365 & 0.6062 & 0.4225 & 0.7073 & 0.6062 & 1.0000 & 0.3949 \\ 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 0.3949 & 1.0000 \end{pmatrix}$$

As $C^{16} = C^8$, then C^8 is an equivalent correlation matrix.

Step 3. For a confidence level λ, according to Eq. (2.24), we can construct a λ-cutting matrix $C_\lambda = \left(\rho_{ij}^{\lambda}\right)_{m \times m}$. Different λ produces different λ-cutting matrix $C_\lambda = \left(\rho_{ij}^{\lambda}\right)_{m \times m}$.

Step 4. Based on the derived λ-cutting matrix $C_\lambda = \left(\rho_{ij}^{\lambda}\right)_{m \times m}$, we can classify these 10 firms $A_j (j = 1, 2, \ldots, 10)$ into different clusters. The possible classifications of these firms are shown in Table 2.6.

Chen et al. (2013a) utilized the correlation coefficient formula in the form of Eq. (2.2) to conduct the cluster analysis and produced a correlation matrix and an equivalent correlation matrix as well, based on which, some clustering results were obtained (see Table 2.7).

Comparing our method with that of Chen et al. (2013a), the superiorities are significant. Firstly, in terms of the correlation matrix, our correlation matrix consists of different values varying from negative values to positive values; however, in Chen et al. (2013a)'s correlation matrix, only positive values can be used, which consequently cannot represent the negative correlation coefficients between the firms. Secondly, as to the equivalent correlation matrix, the value range in $C^{16\prime}$ is from 0.7984 to 1, which is quite narrow, and thus, it may be not quite convincing to distinguish different clusters. But if using our weighted correlation coefficient, the values in the produced equivalent correlation matrix vary from 0.3949 to 1, which is twice wider than that of $C^{16\prime}$, and thus can better reflect the differences between different clusters.

Table 2.6 Clustering results with respect to the correlation coefficient

Class	Confidence level	Clusters
10	$0.9659 < \lambda \leq 1$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
9	$0.9365 < \lambda \leq 0.9659$	$\{A_1\}, \{A_2, A_3\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
8	$0.8665 < \lambda \leq 0.9365$	$\{A_1\}, \{A_2, A_3\}, \{A_4, A_9\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_{10}\}$
7	$0.8272 < \lambda \leq 0.8665$	$\{A_1\}, \{A_2, A_3, A_4, A_9\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_{10}\}$
6	$0.8068 < \lambda \leq 0.8272$	$\{A_1, A_6\}, \{A_2, A_3, A_4, A_9\}, \{A_5\}, \{A_7\}, \{A_8\}, \{A_{10}\}$
5	$0.7073 < \lambda \leq 0.8068$	$\{A_1, A_6\}, \{A_2, A_3, A_4, A_9\}, \{A_5, A_8\}, \{A_7\}, \{A_{10}\}$
4	$0.6062 < \lambda \leq 0.7073$	$\{A_1, A_6\}, \{A_2, A_3, A_4, A_7, A_9\}, \{A_5, A_8\}, \{A_{10}\}$
3	$0.4225 < \lambda \leq 0.6062$	$\{A_1, A_6\}, \{A_2, A_3, A_4, A_5, A_7, A_8, A_9\}, \{A_{10}\}$
2	$0.3949 < \lambda \leq 0.4225$	$\{A_1, A_2, A_3, A_4, A_5, A_6 A_7, A_8, A_9\}, \{A_{10}\}$
1	$0 \leq \lambda \leq 0.3949$	$\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\}$

Table 2.7 Clustering results with respect to Chen et al. (2013a)'s correlation coefficient

Class	Confidence level	Clusters
10	$0.9515 < \lambda \le 1$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
9	$0.9306 < \lambda \le 0.9515$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_5, A_{10}\}$
8	$0.9238 < \lambda \le 0.9306$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4, A_9\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_5, A_{10}\}$
7	$0.9104 < \lambda \le 0.9238$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4, A_7, A_9\}, \{A_6\}, \{A_8\}, \{A_5, A_{10}\}$
6	$0.9025 < \lambda \le 0.9104$	$\{A_1, A_6\}, \{A_2\}, \{A_3\}, \{A_4, A_7, A_9\}, \{A_8\}, \{A_5, A_{10}\}$
5	$0.8997 < \lambda \le 0.9025$	$\{A_1, A_6\}, \{A_2\}, \{A_3\}, \{A_4, A_7, A_8, A_9\}, \{A_5, A_{10}\}$
4	$0.8520 < \lambda \le 0.8997$	$\{A_1, A_6\}, \{A_2\}, \{A_3, A_4, A_7, A_8, A_9\}, \{A_5, A_{10}\}$
3	$0.8200 < \lambda \le 0.8520$	$\{A_1, A_6\}, \{A_2\}, \{A_3, A_4, A_5, A_7, A_8, A_9, A_{10}\}$
2	$0.7984 < \lambda \le 0.8200$	$\{A_1, A_6\}, \{A_2, A_3, A_4, A_5, A_7, A_8, A_9, A_{10}\}$
1	$0 \le \lambda \le 0.7984$	$\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\}$

2.2 Novel Entropy Measures of Hesitant Fuzzy Sets

2.2.1 The Existing Entropy Measures of Hesitant Fuzzy Sets

Motivated by the axiomatic definition of entropy for fuzzy sets, Xu and Xia (2012) proposed the principles of entropy measure for HFE in terms of the fuzziness of a HFE.

Definition 2.7 (*Xu and Xia* 2012). A real-valued function $E : H \to [0, 1]$ is called an entropy for the HFE α, if it satisfies:

(1) $E(\alpha) = 0$ if and only if $\alpha = \{0\}$ or $\alpha = \{1\}$.
(2) $E(\alpha) = 1$ if and only if $\alpha_{\sigma(i)} + \alpha_{\sigma(l-i+1)} = 1$, for $i = 1, 2, \ldots, l_\alpha$.
(3) $E(\alpha) \le E(\beta)$ if $\alpha_{\sigma(i)} \le \beta_{\sigma(i)}$ for $\beta_{\sigma(i)} + \beta_{\sigma(l-i+1)} \le 1$ or $\alpha_{\sigma(i)} \ge \beta_{\sigma(i)}$ for $\beta_{\sigma(i)} + \beta_{\sigma(l-i+1)} \ge 1$, $i = 1, 2, \ldots, l$.
(4) $E(\alpha) = E(\alpha^c)$.

Based on Definition 2.7, Xu and Xia (2012) introduced some entropy measures for a HFE α.

$$E_1(\alpha) = \frac{1}{l_\alpha(\sqrt{2} - 1)} \sum_{i=1}^{l_\alpha} \left(\sin \frac{\pi\left(\alpha_{\sigma(i)} + \alpha_{\sigma(l_\alpha - i + 1)}\right)}{4} + \sin \frac{\pi\left(2 - \alpha_{\sigma(i)} - \alpha_{\sigma(l_\alpha - i + 1)}\right)}{4} - 1 \right) \tag{2.25}$$

$$E_2(\alpha) = \frac{1}{l_\alpha(\sqrt{2} - 1)} \sum_{i=1}^{l_\alpha} \left(\cos \frac{\pi\left(\alpha_{\sigma(i)} + \alpha_{\sigma(l_\alpha - i + 1)}\right)}{4} + \cos \frac{\pi\left(2 - \alpha_{\sigma(i)} - \alpha_{\sigma(l_\alpha - i + 1)}\right)}{4} - 1 \right) \tag{2.26}$$

$$E_3(\alpha) = -\frac{1}{l_\alpha \ln 2}\sum_{i=1}^{l_\alpha}\left(\frac{\alpha_{\sigma(i)}+\alpha_{\sigma(l_\alpha-i+1)}}{2}\ln\frac{\alpha_{\sigma(i)}+\alpha_{\sigma(l_\alpha-i+1)}}{2} + \frac{2-\alpha_{\sigma(i)}-\alpha_{\sigma(l_\alpha-i+1)}}{2}\ln\frac{2-\alpha_{\sigma(i)}-\alpha_{\sigma(l_\alpha-i+1)}}{2}\right) \tag{2.27}$$

$$E_4(\alpha) = -\frac{1}{l_\alpha(2^{(1-s)t}-1)}\sum_{i=1}^{l_\alpha}\left[\left(\left(\frac{\alpha_{\sigma(i)}+\alpha_{\sigma(l_\alpha-i+1)}}{2}\right)^s + \left(\frac{2-\alpha_{\sigma(i)}-\alpha_{\sigma(l_\alpha-i+1)}}{2}\right)^s\right)^t - 1\right]$$

$$t \neq 0,\ s \neq 1,\ s > 0 \tag{2.28}$$

All the above entropy measures satisfy the conditions in Definition 2.7. However, if we apply them to the HFEs whose complements are equal to themselves, we can get the same entropy degree. This indicates that the entropy measures introduced by Xu and Xia (2012) cannot correctly discriminate different HFEs in some cases.

Example 2.5 (*Zhao et al.* 2015). Let $\alpha_1 = \{0.2, 0.5, 0.8\}$ and $\alpha_2 = \{0.4, 0.5, 0.6\}$ be two HFEs. Obviously, $\alpha_1 = \alpha_1^c$, $\alpha_2 = \alpha_2^c$ and the fuzziness of α_2 is greater than that of α_1. Applying the entropy measures E_i $(i = 1, 2, 3, 4)$ to the HFEs α_1 and α_2, we obtain $E_i(\alpha_1) = E_i(\alpha_2) = 1$, for $i = 1, 2, 3, 4$, which are not consistent with our intuition.

Based on the distance measure between HFEs (Xu and Xia 2011a), Farhadinia (2013) gave the following axiomatic definition of entropy to measure the fuzziness of a HFE.

Definition 2.8 (*Farhadinia* 2013). Let $d(\alpha, \{0.5\})$ be the distance between the HFE α and $\{0.5\}$. A real function $E_d : H \to [0, 1]$ is called a distance-based entropy for the HFE α, if it satisfies:

(1) $E_d(\alpha) = 0$ if and only if $\alpha = \{0\}$ or $\alpha = \{1\}$.
(2) $E_d(\alpha) = 1$ if and only if $\alpha = \{0.5\}$.
(3) $E_d(\alpha) \le E_d(\beta)$ if $d(\alpha, \{0.5\}) \ge d(\beta, \{0.5\})$.
(4) $E_d(\alpha) = E_d(\alpha^c)$.

Theorem 2.4 provided an approach to generate the distance-based entropy measures for HFEs.

Theorem 2.4 (Farhadinia 2013). *Let* $Z : [0, 1] \to [0, 1]$ *be a strictly monotone decreasing real function, and* $d(\alpha, \{0.5\})$ *be the distance between the HFE* α *and* $\{0.5\}$. *Then,*

$$E_d(\alpha) = \frac{Z(2d(\alpha, \{0.5\})) - Z(1)}{Z(0) - Z(1)} \tag{2.29}$$

is an entropy measure for the HFE α.

Xu and Xia (2011a) defined three kinds of distance measures which can be used to calculate the distance between the HFE α and $\{0.5\}$.

$$d_{1k}(\alpha, \{0.5\}) = \left[\frac{1}{l}\sum_{i=1}^{l}\left|\alpha_{\sigma(i)} - 0.5\right|^k\right]^{1/k}, k = 1, 2 \tag{2.30}$$

$$d_{2k}(\alpha, \{0.5\}) = \max_i\left\{\left|\alpha_{\sigma(i)} - 0.5\right|^k\right\}, k = 1, 2 \tag{2.31}$$

$$d_{3k}(\alpha, \{0.5\}) = \left\{\left[\frac{1}{l}\sum_{i=1}^{l}\left|\alpha_{\sigma(i)} - 0.5\right|^k\right]^{1/k} + \max_i\left\{\left|\alpha_{\sigma(i)} - 0.5\right|^k\right\}\right\}, k = 1, 2 \tag{2.32}$$

Let α be three HFEs $\{0, 1\}$, $\{0\}$ and $\{1\}$, respectively. Then by Eqs. (2.30)–(2.32), we can calculate

$$d_{1k}(\{0,1\},\{0.5\}) = d_{1k}(\{0\},\{0.5\}) = d_{1k}(\{1\},\{0.5\}) = \frac{1}{2}, k = 1, 2$$
$$d_{2k}(\{0,1\},\{0.5\}) = d_{2k}(\{0\},\{0.5\}) = d_{2k}(\{1\},\{0.5\}) = \left(\frac{1}{2}\right)^k, k = 1, 2$$
$$d_{3k}(\{0,1\},\{0.5\}) = d_{3k}(\{0\},\{0.5\}) = d_{3k}(\{1\},\{0.5\}) = \frac{1}{2}\left[\frac{1}{2} + \left(\frac{1}{2}\right)^k\right], k = 1, 2$$

According to Theorem 2.4, we get

$$E_{d_{1k}}(\{0,1\}) = E_{d_{1k}}(\{0\}) = E_{d_{1k}}(\{1\}) = 0, k = 1, 2$$
$$E_{d_{21}}(\{0,1\}) = E_{d_{21}}(\{0\}) = E_{d_{21}}(\{1\}) = 0$$
$$E_{d_{22}}(\{0\}) = E_{d_{22}}(\{1\}) = \frac{Z(1/2) - Z(1)}{Z(0) - Z(1)} \neq 0$$
$$E_{d_{31}}(\{0,1\}) = E_{d_{31}}(\{0\}) = E_{d_{31}}(\{1\}) = 0$$
$$E_{d_{32}}(\{0\}) = E_{d_{32}}(\{1\}) = \frac{Z(3/8) - Z(1)}{Z(0) - Z(1)} \neq 0$$

The above results reveal that no matter which distance measure we employ, the derived entropies for HFEs do not meet the first condition in Definition 2.8, which implies that the entropy measure in Eq. (2.29) is unreasonable. Moreover, for any two HFEs α and β, if $d(\alpha, \{0.5\}) = d(\beta, \{0.5\})$, then by Eq. (2.29), we have $E(\alpha) = E(\beta)$. That is to say, different HFEs that have the same distance from the HFE $\{0.5\}$ would yield the same entropy in case we use the entropy measure proposed in Theorem 2.4. This is definitely unreasonable.

Particularly, let $Z(t) = 1 - t$ and d be the hesitant normalized Hamming distance d_{11}, then the entropy measure in Eq. (2.29) turns out to be:

$$E_{d_{11}}(\alpha) = 1 - \frac{2}{l_\alpha} \sum_{i=1}^{l_\alpha} \left| \alpha_{\sigma(i)} - \frac{1}{2} \right| \tag{2.33}$$

Example 2.6 (*Zhao et al.* 2015). In a multiple criteria decision making problem, two decision organizations consider the possible membership degrees of x to the set M. The experts in the first organization think that the membership degree should be 0.01, while in the second organization, some experts deem it as 0.01, and the others deem it as 0.99. Then, the membership degree provided by the first organization is 0.01, which is very small, and thus, we can easily deduce that the experts in the first organization are inclined to consider that x does not belong to the set M. Similarly, it can be easily deduced that some experts in the second organization tend to think that x does not belong to the set M, and the others tend to believe that x belongs to the set M, and the degrees that x belongs to and not to the set M are the same, which implies that according to the decision information provided by the second decision organization, we are not sure whether x belongs to the set M or not. Thus, we may say that the decision information offered by the first organization is more specific than that offered by the other one. We can use the HFEs $\alpha_1 = \{0.01\}$ and $\alpha_2 = \{0.01, 0.99\}$ to represent the possible membership degrees of x into M provided by these two organizations, respectively. Then, by Eq. (2.33), we get $E_{d_{11}}(\alpha_1) = E_{d_{11}}(\alpha_2) = 0.02$, which is unreasonable because these two HFEs are significantly different in terms of specificity based on the above analysis.

2.2.2 Novel Two-Tuple Entropy Measures of Hesitant Fuzzy Sets

As mentioned above, the entropy measures proposed by Xu and Xia (2012) and Farhadinia (2013) are incapable to effectively distinguish HFEs in many cases. In our opinion, for a HFE, except for the fuzziness, there exists another kind of uncertainty, i.e., non-specificity. The fuzziness of a HFE is related to the deviation between the HFE and its nearest crisp set, while the non-specificity is related to the imprecise knowledge contained in the HFE. Suppose that the membership degrees of the element x to the set A provided by a decision organization are presented by the HFE $h(x) = \{0, 1\}$. From the HFE $h(x) = \{0, 1\}$, we know that the membership degree of x to A may be 0 indicating that x absolutely does not belong to A, and may be 1 implying that x completely belongs to A. We are not sure whether the element x belongs to the set A or not. That is to say, this case involves non-specificity. Non-specificity is another kind of uncertainty associated with a HFE. In this section, we present a new axiomatic definition of the entropy for HFEs, which captures the two types of uncertainty associated with a HFE. Then, we introduce some methods to construct the entropy measures for HFE.

Definition 2.9 (*Zhao et al*. 2015). Let $E_F, E_{NS} : H \to [0, 1]$ be two real functions. The pair (E_F, E_{NS}) is called a two-tuple entropy measure for the HFE α if E_F satisfies the following axiomatic requirements:

(1) $E_F(\alpha) = 0$ if and only if α is crisp, that is, $\alpha = \{0\}$ or $\alpha = \{1\}$;
(2) $E_F(\alpha) = 1$ if and only if $\alpha = \{0.5\}$;
(3) $E_F(\alpha) = E_F(\alpha^c)$;
(4) For any $i = 1, 2, \ldots, l$, if $\alpha_{\sigma(i)} \leq \beta_{\sigma(i)}$ for $\beta_{\sigma(i)} \leq 0.5$ or if $\alpha_{\sigma(i)} \geq \beta_{\sigma(i)}$ for $\beta_{\sigma(i)} \geq 0.5$, then $E_F(\alpha) \leq E_F(\beta)$, and E_{NS} satisfies the following axiomatic requirements:
(5) $E_{NS}(\alpha) = 0$ if and only if there is only one value in α, that is, $\alpha = \{u\}$ with $0 \leq u \leq 1$;
(6) $E_{NS}(\alpha) = 1$ if and only if $\alpha = \{0, 1\}$;
(7) $E_{NS}(\alpha) = E_{NS}(\alpha^c)$;
(8) $E_{NS}(\alpha) \geq E_{NS}(\beta)$ if for any $i, j = 1, 2, \ldots, l$, $\left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right| \geq \left|\beta_{\sigma(i)} - \beta_{\sigma(j)}\right|$.

Definition 2.9 uses a pair (E_F, E_{NS}) to represent the two kinds of uncertainty linked to a HFE where E_F, called the fuzzy entropy, is considered as a measure of fuzziness to quantify how far the HFE is from its closest crisp set, and E_{NS}, called the non-specific entropy, is proposed to measure the non-specificity of a HFE. It is noticed that the proposed non-specificity measure differs from that linked to the fuzzy set or the IFS. The introduced two-tuple entropy measure (E_F, E_{NS}) not only maintains the traditional properties of entropy, i.e., measuring the fuzziness aspect of uncertainty, but also reflects another aspect of uncertainty, i.e., non-specificity.

(1) **Fuzzy entropy** E_F

In this part, we provide some methods to generate the measures to quantify the fuzziness of a HFE.

Theorem 2.5 (Zhao et al. 2015). *Let* $R : [0, 1]^2 \to [0, 1]$ *be a mapping and satisfy:*

(1) $R(x, y) = 0$ *if and only if* $x = y = 0$ *or* $x = y = 1$.
(2) $R(x, y) = 1$ *if and only if* $x = y = 0.5$.
(3) $R(x, y) = R(1 - y, 1 - x)$ *for all* $x, y \in [0, 1]$.
(4) *If* $0 \leq x_1 \leq x_2 \leq 0.5, 0 \leq y_1 \leq y_2 \leq 0.5$, *then* $R(x_1, y_1) \leq R(x_2, y_2)$; *if* $0.5 \leq x_1 \leq x_2 \leq 1, 0.5 \leq y_1 \leq y_2 \leq 1$, *then* $R(x_1, y_1) \geq R(x_2, y_2)$.

Then the mapping $E_F : H \to [0, 1]$ *defined as*

$$E_F(\alpha) = \frac{2}{l_\alpha(l_\alpha + 1)} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} R\left(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}\right) \tag{2.34}$$

fulfills the axioms (1)–(4) in Definition 2.9.

Proof

(1) If $\alpha = \{0\}$, then by Eq. (2.34), we get $E_F(\{0\}) = R(0,0) = 0$; if $\alpha = \{1\}$, then $E_F(\{1\}) = R(1,1) = 0$. Conversely, if $E_F(\alpha) = 0$, then $R(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}) = 0$, for any $i,j = 1,2,\ldots,l_\alpha$, $j \geq i$. According to the property (1) in Theorem 2.5, we have $\alpha_{\sigma(i)} = 0$ or $\alpha_{\sigma(i)} = 1$, for any $i = 1,2,\ldots,l_\alpha$. Thus, the condition (1) in Definition 2.9 holds.

(2) If $\alpha = \{0.5\}$, then according to Eq. (2.34), we obtain $E_F(\{0.5\}) = R(0.5, 0.5) = 1$. On the contrary, if $E_F(\alpha) = 1$, then by Eq. (2.34), we have $R(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}) = 1$, for any $i,j = 1,2,\ldots,l_\alpha$, $j \geq i$. According to the property (2) in Theorem 2.5, we have $\alpha_{\sigma(i)} = 0.5$, for any $i = 1,2,\ldots,l_\alpha$. Thus, the condition (2) in Definition 2.9 holds.

(3) By Eq. (2.34), we have $E_F(\alpha^c) = \frac{2}{l_\alpha(l_\alpha+1)}\sum_{i=1}^{l_\alpha}\sum_{j\geq i} R\left(\alpha^c_{\sigma(i)}, \alpha^c_{\sigma(j)}\right)$. Since $\alpha^c_{\sigma(i)} = 1 - \alpha_{\sigma(l_\alpha - i + 1)}$, for $i = 1,2,\ldots,l_\alpha$, then, $E_F(\alpha^c) = \frac{2}{l_\alpha(l_\alpha+1)}\sum_{i=1}^{l_\alpha}\sum_{j\geq i}$ $R(1 - \alpha_{\sigma(l_\alpha - i + 1)}, 1 - \alpha_{\sigma(l_\alpha - j + 1)})$. According to the property (3) in Theorem 2.5, we have $E_F(\alpha^c) = \frac{2}{l_\alpha(l_\alpha+1)}\sum_{i=1}^{l_\alpha}\sum_{j\geq i} R(\alpha_{\sigma(l_\alpha - j + 1)}, \alpha_{\sigma(l_\alpha - i + 1)}) = E_F(\alpha)$. Thus, the condition (3) in Definition 2.9 holds.

(4) For any $i = 1,2,\ldots,l$, $\alpha_{\sigma(i)} \leq \beta_{\sigma(i)} \leq 0.5$, according to the property (4) in Theorem 2.5, we get $R(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}) \leq R\left(\beta_{\sigma(i)}, \beta_{\sigma(j)}\right)$, $i,j = 1,2,\ldots,l$, $j \geq i$. Then according to Eq. (2.34), we gain $E_F(\alpha) \leq E_F(\beta)$. The other case can be illustrated in a similar way. Thus, the condition (4) in Definition 2.9 holds.

Remark It is observed that $E_F(\alpha)$ is a fuzzy entropy for the HFE α. By Eq. (2.34), we have $E_F(\{0,1\}) = \frac{1}{3}R(0,1) \neq 0$, which shows that the fuzzy entropy of $\{0,1\}$ is different from those of the HFEs $\{0\}$ and $\{1\}$.

Theorem 2.6 (Zhao et al. 2015). *Let* $\bar{R} : [0,1]^2 \to [0,1]$ *be a mapping and the mapping* $E_F : H \to [0,1]$ *defined as:*

$$E_F(\alpha) = \frac{2}{l_\alpha(l_\alpha+1)}\sum_{i=1}^{l_\alpha}\sum_{j\geq i}\bar{R}\left(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}\right) \tag{2.35}$$

satisfies the axioms (1)–(4) in Definition 2.9, then

(1) $\bar{R}(0,0) = 0$ *and* $\bar{R}(1,1) = 0$.
(2) $\bar{R}(0.5, 0.5) = 1$.
(3) *If for all* $x,y \in [0,1]$, $\bar{R}(x,y) = \bar{R}(y,x)$, *then* $\bar{R}(x,y) = \bar{R}(1-y, 1-x)$ *for all* $x,y \in [0,1]$.
(4) *If* $0 \leq x_1 \leq x_2 \leq 0.5$, *then* $\bar{R}(x_1,x_1) \leq \bar{R}(x_2,x_2)$; *if* $0.5 \leq x_1 \leq x_2 \leq 1$, *then* $\bar{R}(x_1,x_1) \geq \bar{R}(x_2,x_2)$.

Proof (1) and (2) are easy to check, thus, we here only give the proofs of (3) and (4).

(3) Suppose that there exist $x, y \in [0,1]$ such that $\bar{R}(x,y) \neq \bar{R}(1-y, 1-x)$. Without loss of generality, assume that $x \leq y$ and $\bar{R}(x,y) > \bar{R}(1-y, 1-x)$. Given a HFE $\alpha = \{\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\}$ where $\alpha_{\sigma(1)} = x$ and $\alpha_{\sigma(2)} = y$, Then by Eq. (2.35), we get

$$E_F(\alpha) = \frac{1}{3}\left[\bar{R}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(1)}\right) + \bar{R}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right) + \bar{R}\left(\alpha_{\sigma(2)}, \alpha_{\sigma(2)}\right)\right]$$

and

$$E_F(\alpha^c) = \frac{1}{3}\left[\bar{R}\left(1-\alpha_{\sigma(2)}, 1-\alpha_{\sigma(2)}\right) + \bar{R}\left(1-\alpha_{\sigma(2)}, 1-\alpha_{\sigma(1)}\right) + \bar{R}\left(1-\alpha_{\sigma(1)}, 1-\alpha_{\sigma(1)}\right)\right]$$

According to the condition (3) in Definition 2.9, we have $E_F\left(\{\alpha_{\sigma(1)}\}\right) = E_F\left(\{1-\alpha_{\sigma(1)}\}\right)$ and $E_F\left(\{\alpha_{\sigma(2)}\}\right) = E_F\left(\{1-\alpha_{\sigma(2)}\}\right)$, that is, $\bar{R}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(1)}\right) = \bar{R}\left(1-\alpha_{\sigma(1)}, 1-\alpha_{\sigma(1)}\right)$ and $\bar{R}\left(\alpha_{\sigma(2)}, \alpha_{\sigma(2)}\right) = \bar{R}\left(1-\alpha_{\sigma(2)}, 1-\alpha_{\sigma(2)}\right)$. Thus, $E_F(\alpha) > E_F(\alpha^c)$, which contradicts the condition (3) in Definition 2.9. In other words, the property (3) holds.

(4) Assume that there exist $x_1, x_2 \in [0, 0.5]$ with $x_1 \leq x_2$ such that $\bar{R}(x_1, x_1) > \bar{R}(x_2, x_2)$, then by Eq. (2.35), we get $E_F(\{x_1\}) = \bar{R}(x_1, x_1) > \bar{R}(x_2, x_2) = E_F(\{x_2\})$, which contradicts the condition (4) in Definition 2.9. Similarly, the other case can be proven.

It is not easy to look for the bivariate function R in Theorem 2.5. In what follows, we try to reduce it to a univariate function.

Theorem 2.7 (Zhao et al. 2015). *Let* $\varphi : [0,1] \to [0,1]$ *be a mapping and satisfy:*

(1) $\varphi(x) = 0$ *if and only if* $x = 0$.
(2) $\varphi(x) = 1$ *if and only if* $x = 0.75$.
(3) φ *is monotone non-decreasing in* $[0, 0.75)$ *and monotone non-increasing in* $(0.75, 1]$.

Then, the mapping $E_F : H \to [0,1]$ *defined as*

$$E_F(\alpha) = \frac{2}{l_\alpha(l_\alpha+1)} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} \varphi\left(1 - \alpha_{\sigma(i)}\alpha_{\sigma(j)}\right) \cdot \varphi\left(\alpha_{\sigma(i)} - \alpha_{\sigma(i)}\alpha_{\sigma(j)} + \alpha_{\sigma(j)}\right) \quad (2.36)$$

fulfills the axioms (1)–(4) in Definition 2.9.

Proof Suppose that φ is defined as the above statement and $R(x,y) = \varphi(1-xy) \cdot \varphi(x - xy + y)$. Then we only need to prove that $R(x,y)$ possesses the properties (1)–(4) in Theorem 2.5.

(1) If $R(x,y) = 0$, that is, $\varphi(1-xy) \cdot \varphi(x-xy+y) = 0$, then $\varphi(1-xy) = 0$ or $\varphi(x-xy+y) = 0$. If $\varphi(1-xy) = 0$, then by the condition (1) in this theorem, we get $xy = 1$. Thus, $x = y = 1$. If $\varphi(x-xy+y) = 0$, then $x - xy + y = 0$.

Therefore, we deduce that $x = y = 0$. The converse is easy to prove. Accordingly, the property (1) in Theorem 2.5 holds.

(2) If $R(x, y) = 1$, that is, $\varphi(1 - xy) \cdot \varphi(x - xy + y) = 1$, then we get $\varphi(1 - xy) = 1$ and $\varphi(x - xy + y) = 1$. By the condition (2) in this theorem, we deduce that $xy = 0.25$ and $x - xy + y = 0.75$, from which we get $x = y = 0.5$. It is easy to prove the converse. Then we finish the proof of property (2) in Theorem 2.5.

(3) $$\begin{aligned} R(1-y, 1-x) &= \varphi(1-(1-y)(1-x)) \cdot \varphi(1-y-(1-y)(1-x)+1-x) \\ &= \varphi(x-xy+y) \cdot \varphi(1-xy) = R(x, y) \end{aligned}$$

(4) Assume $0 \le x_1 \le x_2 \le 0.5$ and $0 \le y_1 \le y_2 \le 0.5$, then $0.75 \le 1 - x_2y_2 \le 1 - x_1y_1 \le 1$ and $0 \le x_1 - x_1y_1 + y_1 \le x_2 - x_2y_2 + y_2 \le 0.75$. By the condition (3) in this theorem, we obtain $\varphi(1 - x_1y_1) \le \varphi(1 - x_2y_2)$ and $\varphi(x_1 - x_1y_1 + y_1) \le \varphi(x_2 - x_2y_2 + y_2)$, from which we derive $R(x_1, y_1) \le R(x_2, y_2)$. Similarly, the other case can be illustrated. Thus, the property (4) in Theorem 2.5 holds.

Based on Theorem 2.7, we can set out two entropy measures for HFEs as illustrative examples.

(1) Let $\varphi(t) = 1 - \left(\frac{1}{3}|4t - 3|\right)^r$ with $r \ge 1$. Obviously, φ satisfies the conditions in Theorem 2.7. Then we get the following entropy measure for HFEs:

$$\begin{aligned} E_F^r(\alpha) = \frac{2}{l_\alpha(l_\alpha+1)} \sum_{i=1}^{l_\alpha} \sum_{j \ge i} \left(1 - \left(\frac{1}{3}\left|4\alpha_{\sigma(i)}\alpha_{\sigma(j)} - 1\right|\right)^r\right) \cdot \\ \left(1 - \left(\frac{1}{3}\left|4\alpha_{\sigma(i)} - 4\alpha_{\sigma(i)}\alpha_{\sigma(j)} + 4\alpha_{\sigma(j)} - 3\right|\right)^r\right) \end{aligned} \quad (2.37)$$

For the simplicity of calculation, we take $r = 1$. Then the entropy measure in Eq. (2.37) becomes

$$E_F^1(\alpha) = \frac{2}{l_\alpha(l_\alpha+1)} \sum_{i=1}^{l_\alpha} \sum_{j \ge i} \frac{1}{9}\left(3 - \left|4\alpha_{\sigma(i)}\alpha_{\sigma(j)} - 1\right|\right) \cdot \left(3 - \left|4\alpha_{\sigma(i)} - 4\alpha_{\sigma(i)}\alpha_{\sigma(j)} + 4\alpha_{\sigma(j)} - 3\right|\right) \quad (2.38)$$

(2) Let $\varphi(t) = \frac{2}{3}[\min(2t - 1, 2 - 2t) + 1]$. Then φ satisfies the conditions in Theorem 2.7, and the generated entropy measure for HFEs is

$$\begin{aligned} E_F(\alpha) = \frac{2}{l_\alpha(l_\alpha+1)} \sum_{i=1}^{l_\alpha} \sum_{j \ge i} \frac{4}{9}\left(\min\left(1 - 2\alpha_{\sigma(i)}\alpha_{\sigma(j)}, 2\alpha_{\sigma(i)}\alpha_{\sigma(j)}\right) + 1\right) \cdot \\ \left(\min\left(1 - 2\left(1 - \alpha_{\sigma(i)}\right)\left(1 - \alpha_{\sigma(j)}\right), 2\left(1 - \alpha_{\sigma(i)}\right)\left(1 - \alpha_{\sigma(j)}\right)\right) + 1\right) \end{aligned} \quad (2.39)$$

The following example illustrates that the entropy measures proposed in this chapter can produce better results than those introduced by Xu and Xia (2012) in distinguishing HFEs.

Table 2.8 The results obtained by different entropy measures

Results	$E_F^1(\alpha_i)$	$E_F^2(\alpha_i)$	$E_F^3(\alpha_i)$	$E_F(\alpha_i)$
$i = 1$	0.8696	0.9851	0.9981	0.8696
$i = 2$	0.5065	0.7385	0.8391	0.5065

Example 2.7 (*Zhao et al.* 2015). Consider two HFEs $\alpha_1=\{0.4, 0.5, 0.6\}$ and $\alpha_2=\{0.1, 0.5, 0.9\}$. Obviously, $\alpha_1=\alpha_1^c$ and $\alpha_2 = \alpha_2^c$, and intuitively, the fuzziness of α_1 should be greater than that of α_2. Utilizing the entropy measures shown in Eqs. (2.25)–(2.28) to calculate the entropy of the HFE $\alpha_i (i = 1, 2)$, we get $E_j(\alpha_1) = E_j(\alpha_2) = 1\ (j = 1, 2, 3, 4)$, which are counter-intuitive. On the contrary, if we use the entropy measures shown in Eqs. (2.37)–(2.39), we can get different results presented in Table 2.8.

From Table 2.8, we can find that no matter which entropy measures we use, the entropy of α_1 is always greater than that of α_2. This is consistent with our intuition. In other words, the proposed entropy measures are able to overcome the drawback of Xu and Xia (2012)'s entropy measures, that is, those measures cannot differentiate the different HFEs which are equal to their complements.

(2) **Non-specific entropy** E_{NS}

Now we pay attention to the other aspect of uncertainty associated with a HFE, i.e., the non-specificity, and introduce some measures to quantify the non-specificity of a HFE.

Let

$$\langle l_\alpha \rangle = \begin{cases} 2, & l_\alpha = 1 \\ l_\alpha(l_\alpha - 1), & l_\alpha \geq 2 \end{cases}$$

Firstly, we give the following general result:

Theorem 2.8 (Zhao et al. 2015). *Let* $F : [0,1]^2 \rightarrow [0,1]$ *be a mapping and satisfy:*

(1) $F(x, y) = 0$ *if and only if* $x = y$.
(2) $F(x, y) = 1$ *if and only if* $\{0, 1\} \cap \{x, y\} \neq \phi$.
(3) $F(x, y) = F(1 - y, 1 - x)$ *for all* $x, y \in [0, 1]$.
(4) *For* $x, y, z, w \in [0, 1]$, *if* $|x - y| \geq |z - w|$, *then* $F(x, y) \geq F(z, w)$.

Then the mapping $E_{NS} : H \rightarrow [0, 1]$ *defined as:*

$$E_{NS}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} F\left(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}\right) \tag{2.40}$$

satisfies the axioms (5)–(8) in Definition 2.9.

Proof

(1) If there is only one value in the HFE α, that is, $\alpha = \{u\}$, then by Eq. (2.40), we get $E_{NS}(\alpha) = F(u, u) = 0$.
Conversely, if $E_{NS}(\alpha) = 0$, then $F(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}) = 0$, for any $i, j = 1, 2, \ldots, l_\alpha, j \geq i$. According to the property (1) in this theorem, we deduce that $\alpha_{\sigma(i)} = \alpha_{\sigma(j)}$, for any $i, j = 1, 2, \ldots, l_\alpha, j \geq i$. That is to say, the HFE α has only one value. Thus, the condition (5) in Definition 2.9 holds.

(2) If $\alpha = \{0, 1\}$, then according to Eq. (2.40) and the property (1) and property (2) in this theorem, we have $E_{NS}(\alpha) = F(0,0) + F(0,1) + F(1,1) = 1$.
On the contrary, if $E_{NS}(\alpha) = 1$ with $\alpha = \{\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \ldots, \alpha_{\sigma(l_\alpha)}\}$, $l_\alpha \geq 2$, then by Eq. (2.40), we obtain $F(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}) = 1$, for any $i, j = 1, 2, \ldots, l_\alpha, j > i$. If $l_\alpha = 2$, then according to the property (2) in this theorem, we obtain $\alpha_{\sigma(1)} = 0$ and $\alpha_{\sigma(2)} = 1$, that is, $\alpha = \{0, 1\}$. If $l_\alpha > 2$, for instance, let $l_\alpha = 3$, then we get $F(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}) = 1$, $F(\alpha_{\sigma(1)}, \alpha_{\sigma(3)}) = 1$ and $F(\alpha_{\sigma(2)}, \alpha_{\sigma(3)}) = 1$. By the first two equations, we deduce that $\alpha_{\sigma(1)} = 0$, $\alpha_{\sigma(2)} = 1$ and $\alpha_{\sigma(3)} = 1$, and by the third equation, we deduce that $\alpha_{\sigma(2)} = 0$ and $\alpha_{\sigma(3)} = 1$, which is contradictory. In a similar way, we can illustrate that it is contradictory when l_α takes any value larger than 3. Thus, the condition (6) in Definition 2.9 holds.

(3) The proof of the condition (7) is similar to that of the condition (3) in Theorem 2.5.

(4) The proof of the condition (8) is straightforward according to the property (4) in this theorem.

Theorem 2.9 (Zhao et al. 2015). *Let* $\bar{F} : [0,1]^2 \to [0,1]$ *be a mapping and let the mapping* $E_{NS} : H \to [0,1]$ *defined as:*

$$E_{NS}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} \bar{F}(\alpha_{\sigma(i)}, \alpha_{\sigma(j)}) \tag{2.41}$$

satisfy the axioms (5)–(8) in Definition 2.9*, then*

(1) $\bar{F}(x, y) = 0$ *if and only if* $x = y$.

(2) *If* $\bar{F}(x, y) = \bar{F}(y, x)$*, for all* $x, y \in [0, 1]$*, then (i)* $\bar{F}(x, y) = 1$ *if and only if* $\{0, 1\} \cap \{x, y\} \neq \phi$*; (ii)* $\bar{F}(x, y) = \bar{F}(1-y, 1-x)$ *for all* $x, y \in [0, 1]$*; (iii)* $\bar{F}(x, y) \geq \bar{F}(z, w)$ *if* $|x - y| \geq |z - w|$ *for* $x, y, z, w \in [0, 1]$.

Proof

(1) Assume that there exist $x, y \in [0, 1]$ with $x \neq y$ such that $\bar{F}(x, y) = 0$. Without loss of generality, suppose $x < y$. Consider the HFEs $\beta = \{\alpha_{\sigma(1)}\}$ and $\gamma = \{\alpha_{\sigma(2)}\}$ assigned by $\alpha_{\sigma(1)} = x$ and $\alpha_{\sigma(2)} = y$, respectively. Then according to the requirement (5) in Definition 2.9, we have

$$E_{NS}(\beta) = \bar{F}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(1)}\right) = 0,\ E_{NS}(\gamma) = \bar{F}\left(\alpha_{\sigma(2)}, \alpha_{\sigma(2)}\right) = 0$$

For the HFE $\alpha = \left\{\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right\}$, we have

$$E_{NS}(\alpha) = \bar{F}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(1)}\right) + \bar{F}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right) + \bar{F}\left(\alpha_{\sigma(2)}, \alpha_{\sigma(2)}\right) = 0 \tag{2.42}$$

By the requirement (5) in Definition 2.9, Eq. (2.42) holds if and only if $\alpha_{\sigma(1)}=\alpha_{\sigma(2)}$, that is, $x = y$, which is contradictory. Similarly, the converse can be proven.

(2) (i) Let $\alpha = \{0, 1\}$, then according to Eq. (2.41) and the requirement (6) in Definition 2.9, we get $E_{NS}(\alpha) = \bar{F}(0,0) + \bar{F}(0,1) + \bar{F}(1,1) = 1$. Since $\bar{F}(0,0) = \bar{F}(1,1) = 0$, then we obtain $\bar{F}(0,1) = 1$. Since $\bar{F}(x,y) = \bar{F}(y,x)$, for all $x, y \in [0,1]$, then $\bar{F}(1,0) = 1$. Conversely, suppose that there exist $x, y \in [0,1]$ with $\{0,1\} \cap \{x,y\} = \phi$ such that $\bar{F}(x,y) = 1$. Without loss of generality, assume that $x > y$. Let a HFE α be $\alpha = \left\{\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right\}$ defined by $\alpha_{\sigma(1)} = y$ and $\alpha_{\sigma(2)} = x$. Then by Eq. (2.41), we get

$$E_{NS}(\alpha) = \bar{F}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right) = \bar{F}(y,x) = 1 \tag{2.43}$$

According to the requirement (6) in Definition 2.9, Eq. (2.43) holds if and only if $\alpha_{\sigma(1)} = 0$ and $\alpha_{\sigma(2)} = 1$, which is contradictory.

(ii) Suppose that there exist $x, y \in [0,1]$ such that $\bar{F}(x,y) \neq \bar{F}(1-y, 1-x)$. Without loss of generality, assume that $x \leq y$ and $\bar{F}(x,y) > \bar{F}(1-y, 1-x)$. Given a HFE α as $\alpha = \left\{\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right\}$, where we assign $\alpha_{\sigma(1)} = x$ and $\alpha_{\sigma(2)} = y$, then we get

$$E_{NS}(\alpha) = \bar{F}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right) > \bar{F}\left(1-\alpha_{\sigma(2)}, 1-\alpha_{\sigma(1)}\right) = E_{NS}(\alpha^c)$$

which contradicts the axiomatic requirement (7) in Definition 2.9.

(iii) Suppose that there exist $x, y, z, w \in [0,\ 1]$ with $|x-y| \geq |z-w|$ such that $\bar{F}(x,y) < \bar{F}(z,w)$. Without loss of generality, assume that $x \leq y$ and $z \leq w$. Considering the HFE $\alpha = \left\{\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right\}$ defined as $\alpha_{\sigma(1)} = x$ and $\alpha_{\sigma(2)} = y$, and the HFE $\beta = \left\{\beta_{\sigma(1)}, \beta_{\sigma(2)}\right\}$ given by $\beta_{\sigma(1)} = z$ and $\beta_{\sigma(2)} = w$, we obtain

$$E_{NS}(\alpha) = \bar{F}\left(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}\right) < \bar{F}\left(\beta_{\sigma(1)}, \beta_{\sigma(2)}\right) = E_{NS}(\beta)$$

which contradicts the requirement (8) in Definition 2.9. This completes the proof of Theorem 2.8.

Bustince et al. (2012) introduced the grouping function to measure to what extent an element belongs to at least one of two given classes.

Definition 2.10 (*Bustince et al.* 2012). A grouping function is a mapping $G : [0,1]^2 \rightarrow [0,1]$ such that:

(1) $G(x, y) = G(y, x)$ for all $x, y \in [0,1]$.
(2) $G(x, y) = 0$ if and only if $x = y = 0$.
(3) $G(x, y) = 1$ if and only if $x = 1$ or $y = 1$.
(4) G is monotonically increasing in both variables.

We can construct the non-specific entropy measure for HFE by means of the grouping function.

Theorem 2.10 (Zhao et al. 2015). *Let G be a grouping function. Then the mapping* $E_{NSG} : H \rightarrow [0,\ 1]$ *shown as:*

$$E_{FBG}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} G\left(\left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right|, \left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right|\right)$$

defines a non-specific entropy measure for the HFE α *satisfying the axioms (5)–(8) in Definition* 2.9.

Proof It is observed that the mapping $F(x, y) = G(|x - y|, |x - y|)$ satisfies the properties (1)–(4) stated in Theorem 2.8. Thus, $E_{FBG}(\alpha)$ is a non-specific entropy measure for α.

If we define $G : [0,1]^2 \rightarrow [0,1]$ as $G(x, y) = x + y - xy$, then G satisfies the conditions in Definition 2.10. That is to say, G is a grouping function. Thus, based on Theorem 2.10, we get a non-specific entropy measure:

$$E^1_{NSG}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} 2\left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right| - \left(\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right)^2 \tag{2.44}$$

If we define $G : [0,1]^2 \rightarrow [0,1]$ as $G(x, y) = 1 - \frac{\sqrt{(1-x)(1-y)}}{\sqrt{(1-x)(1-y) + 1 - (1-x)(1-y)}}$, then G is a grouping function, and the corresponding non-specific entropy measure is

$$E^2_{NSG}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} 1 - \frac{1 - \left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right|}{1 + \left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right| - \left(\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right)^2} \tag{2.45}$$

Clearly, it is a bit difficult to look for such a bivariate function satisfying the conditions in Theorem 2.8. Below we attempt to reduce it to a univariate function.

Theorem 2.11 (Zhao et al. 2015). *Let* $g : [0,1] \rightarrow [0,1]$ *be a mapping and satisfy:*

(1) $g(x) = 0$ *if and only if* $x = 0$.
(2) $g(x) = 1$ *if and only if* $x = 1$.
(3) g *is monotone non-decreasing.*

Then the mapping $E_{NS} : H \rightarrow [0,1]$ *defined as:*

$$E_{NS}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} g\left(\left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right|\right) \tag{2.46}$$

satisfies the axioms (5)–(8) in Definition 2.9.

Proof Let $F(x,y) = g(|x-y|)$, then the mapping $F(x,y)$ satisfies the properties in Theorem 2.8.

Below we give several specific examples to illustrate Theorem 2.11.

(1) Let $g : [0,1] \to [0,1]$ be defined as $g(t) = \frac{2t}{1+t}$. It satisfies the conditions in Theorem 2.11. Thus, the corresponding non-specific entropy measure is

$$E_{NS}^1(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} \frac{2\left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right|}{\left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right| + 1} \tag{2.47}$$

(2) Let $g : [0,1] \to [0,1]$ be $g(t) = \frac{\lg(1+t)}{\lg 2}$, then g satisfies the conditions in Theorem 2.11, and the corresponding non-specific entropy measure is

$$E_{NS}^2(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} \frac{\lg\left(1 + \left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right|\right)}{\lg 2} \tag{2.48}$$

(3) Let $g : [0,1] \to [0,1]$ be $g(t) = te^{t-1}$, then g satisfies the conditions in Theorem 2.11, and the corresponding non-specific entropy measure is

$$E_{NS}^3(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} \left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right| e^{\left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right| - 1} \tag{2.49}$$

It can be easily observed that the automorphisms of the unit interval satisfy the conditions in Theorem 2.11. In the following, we set out several non-specific entropy measures produced by them.

(4) Let $\phi : [0,1] \to [0,1]$ be defined as $\phi(t) = t^r$ with $r > 0$. Then ϕ is an automorphism of the unit interval, i.e., ϕ is continuous, strictly increasing and satisfies the conditions $\phi(0) = 0, \ \ \phi(1) = 1$ (Bustince et al. 2003). According to Theorem 2.11, we obtain the corresponding non-specific entropy measure as:

$$E_{NSA}^1(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} \left|\alpha_{\sigma(i)} - \alpha_{\sigma(j)}\right|^r \tag{2.50}$$

Especially, if we take $r = 1$, then the non-specific entropy measure becomes

Table 2.9 The results generated by different non-specific entropy formulas

Result	$E^1_{NSG}(\alpha_i)$	$E^2_{NSG}(\alpha_i)$	$E^1_{NS}(\alpha_i)$	$E^2_{NS}(\alpha_i)$	$E^3_{NS}(\alpha_i)$	$E^2_{NSA}(\alpha_i)$	$E^4_{NSA}(\alpha_i)$
$i = 1$	0.84	0.6774	0.75	0.6781	0.4022	0.6	0.84
$i = 2$	0.2467	0.2196	0.2323	0.1793	0.0571	0.1333	0.2467

$$E^2_{NSA}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} \left| \alpha_{\sigma(i)} - \alpha_{\sigma(j)} \right| \tag{2.51}$$

(5) Let $\phi : [0, 1] \rightarrow [0, 1]$ be defined as $\phi(t) = 1 - (1 - t)^r$ with $r > 0$. Then ϕ is an automorphism of the unit interval. Based on Theorem 2.11, we get the generated non-specific entropy measure:

$$E^3_{NSA}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} 1 - \left(1 - \left| \alpha_{\sigma(i)} - \alpha_{\sigma(j)} \right| \right)^r \tag{2.52}$$

In particular, when $r = 2$, the non-specific entropy measure becomes

$$E^4_{NSA}(\alpha) = \frac{2}{\langle l_\alpha \rangle} \sum_{i=1}^{l_\alpha} \sum_{j \geq i} 1 - \left(1 - \left| \alpha_{\sigma(i)} - \alpha_{\sigma(j)} \right| \right)^2 \tag{2.53}$$

It is noted that the entropy measures introduced by Farhadinia (2013) cannot discriminate the HFEs having the same distance from the HFE $\{0.5\}$. The following example shows that our entropy measures can overcome this drawback perfectly.

Example 2.8 (*Zhao et al.* 2015). Suppose two HFEs $\alpha_1 = \{0.2, 0.8\}$ and $\alpha_2 = \{0.1, 0.2, 0.3\}$. Clearly, the information expressed by α_2 is more specific than that of α_1. Nevertheless, by Eq. (2.33), we get $E_{d_{hnh}}(\alpha_1) = E_{d_{hnh}}(\alpha_2) = 0.4$, which is unreasonable. For α_1 and α_2, applying the proposed non-specific entropy measures (2.47)–(2.53), we can get different results, which are listed in Table 2.9.

From Table 2.9, it can be observed that no matter which measure is applied, we always get that the non-specificity of α_1 is greater than that of α_2, which is consistent with our intuition. From this example, we can see that our non-specific entropy measures can distinguish those HFEs that have the same distance from the HFE $\{0.5\}$, while the entropy measure in Eq. (2.33) cannot.

References

Burillo P, Bustince H (1996) Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. Fuzzy Sets Syst 78:305–316

Bustince H, Burillo P, Soria F (2003) Automorphisms, negations and implication operators. Fuzzy Sets Syst 134:209–229

Bustince H, Pagola M, Mesiar R, Hullermeier E, Herrera F (2012) Grouping, overlap, and generalized bientropic functions for fuzzy modeling of pairwise comparisons. IEEE Trans Fuzzy Syst 20:405–415

Chaudhuri BB, Bhattacharya A (2001) On correlation between two fuzzy sets. Fuzzy Sets Syst 118:447–456

Chen N, Xu ZS, Xia MM (2013) Correlation coefficients of hesitant fuzzy sets and their application to clustering analysis. Appl Math Model 37:2197–2211

Chiang DA, Lin NP (1999) Correlation of fuzzy sets. Fuzzy Sets Syst 102:221–226

De Luca A, Termini S (1972) A definition of a non probabilistic entropy in the setting of fuzzy sets theory. Inf Control 20:301–312

Farhadinia B (2013) Information measures for hesitant fuzzy sets and interval-valued hesitant fuzzy sets. Inf Sci 240:129–144

Friedman M, Last M, Makover Y, Kandel A (2007) Anomaly detection in web documents using crisp and fuzzy-based clustering methodology. Inf Sci 177:467–475

Garmendia L, Yager RR, Trillas E, Salvador A (2003) On t-norms based measures of specificity. Fuzzy Sets Syst 133:237–248

Gerstenkorn T, Manko J (1991) Correlation of intuitionistic fuzzy sets. Fuzzy Sets Syst 44:39–43

Hong DH (2006) Fuzzy measures for a correlation coefficient of fuzzy numbers under T_w (the weakest t-norm)-based fuzzy arithmetic operations. Inf Sci 176:150–160

Hong DH, Hwang SY (1995) Correlation of intuitionistic fuzzy sets in probability spaces. Fuzzy Sets Syst 75:77–81

Hung WL (2001) Using statistical viewpoint in developing correlation of intuitionistic fuzzy sets. Int J Uncertainty Fuzziness Knowl Based Syst 9:509–516

Hung WL, Wu JW (2002) Correlation of intuitionistic fuzzy sets by centroid method. Inf Sci 144:219–225

Kaufmann A, Swanson D (1975) Introduction to the theory of fuzzy subsets: fundamental theoretical elements. Academic Press, New York

Liao HC, Xu ZS, Zeng XJ (2015) Novel correlation coefficients between hesitant fuzzy sets and their application in decision making. Knowl-Based Syst 82:115–127

Liu ST, Gao C (2002) Fuzzy measures for correlation coefficient of fuzzy numbers. Fuzzy Sets Syst 128:267–275

Mitchell HB (2004) A correlation coefficient for intuitionistic fuzzy sets. Int J Intell Syst 19: 483–490

Murithy CA, Pal SK, Dutta-Majumder D (1985) Correlation between two fuzzy membership functions. Fuzzy Sets Syst 17:23–38

Pal NR, Bustince H, Pagola M, Mukherjee UK, Goswami DP, Beliakov G (2013) Uncertainties with Atanassov's intuitionistic fuzzy sets: Fuzziness and lack of knowledge. Inf Sci 228:61–74

Parkash O, Sharma P, Mahajan R (2008) New measures of weighted fuzzy entropy and their applications for the study of maximum weighted fuzzy entropy principle. Inf Sci 178: 2389–2395

Szmidt E, Kacprzyk J (2001) Entropy for intuitionistic fuzzy sets. Fuzzy Sets Syst 118:467–477

Szmidt E, Kacprzyk J (2010) Correlation of intuitionistic fuzzy sets. In: Hüllermeier E, Kruse R, Hoffmann F. (eds), IPMU 2010, LNAI 6178, pp. 169–177

Wei CP, Liang X, Zhang YZ (2012) A comparative analysis and improvement of entropy measures for intuitionistic fuzzy sets. J Syst Sci Math Sci 32:1437–1448

Xu ZS (2006) On correlation measures of intuitionistic fuzzy sets. Lect Notes Comput Sci 4224: 16–24

Xu ZS, Xia MM (2011a) On distance and correlation measures of hesitant fuzzy information. Int J Intell Syst 26:410–425

Xu ZS, Xia MM (2011b) Distance and similarity measures for hesitant fuzzy sets. Inf Sci 181:2128–2138

Xu ZS, Xia MM (2012) Hesitant fuzzy entropy and cross-entropy and their use in multiattribute decision-making. Int J Intell Syst 27(9):799–822

Yager RR (1979) On the measure of fuzziness and negation part I: membership in the unit interval. Int J Gen Syst 5:221–229

Yager RR (1992) Default knowledge and measures of specificity. Inf Sci 61:1–44

Yager RR (1998) Measures of specificity. In: Computational intelligence: soft computing and fuzzy-neuro integration with applications. Springer, pp 94–113

Yager RR (2008) Measures of specificity over continuous spaces under similarity relations. Fuzzy Sets Syst 159:2193–2210

Yu CH (1993) Correlation of fuzzy numbers. Fuzzy Sets Syst 55:303–307

Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353

Zhao N, Xu ZS, Liu FJ (2015) Uncertainty measures for hesitant fuzzy information. Int J Intell Syst 30(7):818–836

http://www.springer.com/978-981-10-3264-6

Hesitant Fuzzy Decision Making Methodologies and Applications
Liao, H.; Xu, Z.
2017, XIII, 275 p. 19 illus., 8 illus. in color., Hardcover
ISBN: 978-981-10-3264-6