# CHAPTER 3

# Relationships Among Several Fuzzy Measures

Yingfang Li, Xingxing He and Keyun Qin

In fuzzy set theory, similarity measure, divergence measure, subsethood measure and fuzzy entropy are four basic concepts. They surface in many fields, such as image processing, fuzzy neural networks, fuzzy reasoning, fuzzy control, and so on. The similarity measure describes the degree of similarity of fuzzy sets $A$ and $B$. The divergence measure describes the degree of difference of fuzzy sets $A$ and $B$. The subsethood measure (also called inclusion measure) is a relation between fuzzy sets $A$ and $B$, which indicates the degree to which $A$ is contained in $B$. The entropy of a fuzzy set is the fuzziness of that set.

Yingfang Li
School of Economics Information Engineering
Southwestern University of Finance and Economics
Chengdu 611130, Sichuan, China
e-mail: liyf@swufe.edu.cn

Xingxing He, Keyun Qin
School of Mathematics
Southwest Jiaotong University
Chengdu 610031, Sichuan, China
e-mail: {x.he, qinkeyun}@home.swjtu.edu.cn

Editors: G.A. Papakostas, A.G. Hatzimichailidis, V.G. Kaburlasos, *Handbook of Fuzzy Sets Comparison - Theory, Algorithms and Applications*
DOI: 10.15579/gcsr.vol6.ch3, GCSR Vol. 6, pp. 43-78, 2016

Table 3.1: Several $t$-norms.

| **Name** | **Symbols** | **Formulae** |
|---|---|---|
| Lukasiewicz | $T_L$ | $\max(x+y-1,0)$ |
| Minimum | $T_M$ | $\min(x,y)$ |
| Product | $T_P$ | $x \cdot y$ |

Table 3.2: Several $t$-conorms.

| **Name** | **Symbols** | **Formulae** |
|---|---|---|
| Lukasiewicz | $S_L$ | $\min(x+y,1)$ |
| Maximum | $S_M$ | $\max(x,y)$ |
| Probabilistic sum | $S_P$ | $x+y-x \cdot y$ |

This chapter focuses on discussing relationships among these four fuzzy measures. All of the fuzzy measures are discussed on discrete universes here; the cases for continuous universes can be researched similarly.

***Keywords*** - Fuzzy measures, similarity measure, divergence measure, subsethood measure, fuzzy entropy.

## 3.1 Generating Fuzzy Equivalencies

In fuzzy logic, to extend connectives of crisp sets to the fuzzy case, there may be multiple alternative ways. The connectives of conjunction, disjunction and negation can be formulated by means of the so-called $t$-norms, $t$-conorms and fuzzy negations, respectively. Let us recall the concepts of some fuzzy connectives [21, 36].

**Definition 1.** An associative, commutative and increasing function $T : [0,1]^2 \to [0,1]$ is called a *t-norm* if $T(x,1) = x$ for all $x \in [0,1]$. An associative, commutative and increasing function $S : [0,1]^2 \to [0,1]$ is called a *t-conorm* if $S(x,0) = x$ for all $x \in [0,1]$.

**Example 1.** Tables 3.1 and 3.2 list several $t$-norms and $t$-conorms, respectively.

**Definition 2.** A function $n : [0,1] \to [0,1]$ is called a *fuzzy negation* if it satisfies:

(n1) $n(0) = 1$ and $n(1) = 0$,

(n2) if $x \leq y$, then $n(x) \geq n(y)$.

A fuzzy negation is said to be *involutive* if

(n3) $(n(n(x))) = x$ for all $x \in [0,1]$.

Fuzzy negations that are involutive are called *strong fuzzy negations*.

**Definition 3.** A continuous, strictly increasing function $\varphi : [x, y] \to [x, y]$ with boundary conditions $\varphi(x) = x, \varphi(y) = y$ is called an *automorphism* of the interval $[x, y]$.

**Definition 4.** A *t*-norm $T$ is left-continuous if $T(x, \underset{j\in J}{\vee} y_j) = \underset{j\in J}{\vee} T(x, y_j)$ holds for an arbitrary index set $J$ and for each $x \in [0,1]$ and $(y_j)_{j\in J} \in [0,\ 1]^J$.

**Definition 5.** The *residuation* of a *t*-norm $T$ is the function $\vec{T} : [0,1]^2 \to [0,\ 1]$ defined for all $x, y \in [0,1]$ by $\vec{T}(x,y) = \sup\{\alpha \in [0,1]\,|T(x,\alpha) \le y\}$.

**Lemma 1.** *Given a left-continuous t-norm $T$, we have*

(1) $\vec{T}(x,y)$ is decreasing with respect to the first variable and increasing with respect to the second variable,

(2) $\vec{T}(1,x) = x$ for all $x \in [0,1]$,

(3) $x \le y$ implies $\vec{T}(x,y) = 1$ for all $x, y \in [0,1]$.

**Definition 6.** The *biresiduation* of a *t*-norm $T$ is the function $\tilde{T} : [0,1]^2 \to [0,1]$ defined for all $x, y \in [0,1]$ by $\tilde{T}(x,y) = \min(\vec{T}(x,y), \vec{T}(y,x))$. It is easy to prove that $\tilde{T}$ satisfies the following properties:

($\tilde{T}1$) $\tilde{T}(x,y) = \tilde{T}(y,x)$ for all $x, y \in [0,1]$,

($\tilde{T}2$) $\tilde{T}(x,x) = 1$ for all $x \in [0,1]$,

($\tilde{T}3$) $\tilde{T}(x,1) = x$ for all $x \in [0,1]$,

($\tilde{T}4$) for all $x, y, z \in [0,1]$, $\min(\tilde{T}(x,y), \tilde{T}(y,z)) \ge \tilde{T}(x,z)$ whenever $x \le y \le z$.

Aggregation operators (also called aggregation functions) are functions with special properties [22]. In this chapter we only consider aggregation functions that take real arguments from the unit interval and produce a real value in the unit interval.

**Definition 7.** A function $M : [0,1]^n \to [0,1]$ is an *aggregation operator* if it satisfies the following properties:

(M1) $\mathrm{M}(\underbrace{0, 0, \ldots, 0}_{ntimes}) = 0$,

(M2) $\mathrm{M}(1, 1, \ldots, 1) = 1$,

(M3) $\mathrm{M}(x, x, \ldots, x) = x$ for all $x \in [0,1]$,

(M4) $M$ is monotonically increasing in all of its arguments.

**Example 2.** As examples of aggregation operators we can take:

(1) The arithmetic mean: $\mathrm{M}(x_1, x_2, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$.

(2) The convex linear combinations: $\mathrm{M}(x_1, x_2, \ldots, x_n) = \lambda \min(x_1, x_2, \ldots, x_n) + (1-\lambda)\max(x_1, x_2, \ldots, x_n)$with $\lambda \in [0,1]$.

(3) $\mathrm{M}(x_1, x_2, \ldots, x_n) = \frac{\max(x_1,x_2,\ldots,x_n)}{\max(x_1,x_2,\ldots,x_n)+\max(1-x_1,1-x_2,\ldots,1-x_n)}$.

Fuzzy equivalence (also called equivalence function) is a fuzzy connective, which has been used to formulate the equivalence of two fuzzy sets. In [18], Fodor and Roubens defined equivalence as a binary operator on the unit interval. In [41], the authors briefly introduced the fuzzy equivalence and proposed three examples of fuzzy equivalencies. Fodor and Roubens [18] defined fuzzy equivalence as a binary function on the unit interval in the following way.

**Definition 8.** A binary function $E : [0,1]^2 \to [0,1]$ is called a *fuzzy equivalence* if it satisfies the following properties:

(E1) $E(x,y) = E(y,x)$ for all $x, y \in [0,1]$,

(E2) $E(x,x) = 1$ for all $x \in [0,1]$,

(E3) $E(1,0) = 0$,

(E4) for all $x, y, z \in [0,1]$, if $x \le y \le z$, then $\min(E(x,y), E(y,z)) \ge E(x,z)$.

Some potential properties of fuzzy equivalencies could be considered [29]:

(E5) $E(x,y) = 1$ if and only if $x = y$ for all $x, y \in [0,1]$,

(E6) $E(x,y) = 0$ if and only if $\min(x,y) = 0$ and $\max(x,y) \neq 0$ for all $x, y \in [0,1]$,

(E7) $E(x,1) = x$ for all $x \in [0,1]$,

(E8) $E(x,y) = E(1-x, 1-y)$ for all $x, y \in [0,1]$,

(E9) $E(x, 1-x) = 0$ if and only if $x = 0$ or $x = 1$.

In [6], Bustince et al. presented the concept of restricted equivalence function, which arises from the concept of fuzzy equivalence and some properties usually demanded from the measures used for comparing images [7, 8]. The conditions for restricted equivalence functions are stronger than those of fuzzy equivalencies.

**Definition 9.** A function $REF : [0,1]^2 \to [0,1]$ is called a *restricted equivalence function* if it satisfies the following properties:

(1) $REF(x,y) = REF(y,x)$ for all $x, y \in [0,1]$,

(2) $REF(x,y) = 1$ if and only if $x = y$ for all $x, y \in [0,1]$,

(3) $REF(x,y) = 0$ if and only if $x = 1$ and $y = 0$ or $x = 0$ and $y = 1$,

(4) $REF(x,y) = REF(n(x), n(y))$ for all $x, y \in [0,1]$, $n$ being a strong fuzzy negation,

(5) for all $x, y, z \in [0,1]$, if $x \le y \le z$, then $\min(REF(x,y), REF(y,z)) \ge REF(x,z)$.

Table 3.3: Several $t$-norms with their residuations and biresiduations.

| $T$ | $\vec{T}(x,y)$ | $E_T(x,y)$ |
|---|---|---|
| $T_L$ | $\vec{T}_L(x,y)=\begin{cases}1-x+y, & x>y,\\ 1, & x\le y.\end{cases}$ | $E_{T_L}(x,y)=1-\lvert x-y\rvert$ |
| $T_M$ | $\vec{T}_M(x,y)=\begin{cases}y, & x>y,\\ 1, & x\le y.\end{cases}$ | $E_{T_M}(x,y)=\begin{cases}\min(x,y), & x\neq y,\\ 1, & x=y.\end{cases}$ |
| $T_P$ | $\vec{T}_P(x,y)=\min(1,\frac{y}{x})$ | $E_{T_P}(x,y)=\frac{\min(x,y)}{\max(x,y)}$ |

One of the most interesting issues related to fuzzy equivalencies is their generation [29]. It is shown that the biresiduation of a t-norm is indeed a fuzzy equivalence by $\tilde{T}1-\tilde{T}4$. By the composition of biresiduations and other fuzzy connectives, we obtain several new fuzzy equivalencies. Let us begin with the fuzzy equivalencies based on biresiduation.

**Proposition 1.** *The biresiduation of a t-norm $T$ is indeed a fuzzy equivalence.*

*Remark* 1. (1) The biresiduation associated to a $t$-norm $T$ is also notated by $E_T$. This will be the notation used in this chapter in order to stress the fact that the biresiduation is a fuzzy equivalence. (2) As it is proved in [36], the fuzzy equivalence $E_T$ based on the biresiduation of a $t$-norm $T$ is also a $T$-indistinguishability operator on the unit interval. Therefore, property E4 is reduced to the inequality: for all $x\le y\le z$, $T(E_T(x,y),E_T(y,z))\le E_T(x,z)\le \min(E_T(x,y),E_T(y,z))$.

**Example 3.** Table 3.3 lists three popular $t$-norms with their residuations and biresiduations. In order to avoid the denominator being zero, we set $\frac{0}{0}=1$.

By the composition of biresiduations and other fuzzy connectives, we obtain several new fuzzy equivalencies.

**Proposition 2.** *Let $\varphi$ be an automorphism of the unit interval and $E$ be a fuzzy equivalence. The function$E_\varphi:[0,1]^2\to[0,1]$ defined for all $x,y\in[0,1]$ by*

$$E_\varphi(x,y)=E(\varphi(x),\varphi(y))$$

*is a fuzzy equivalence.*

**Example 4.** For all $x\in[0,1]$, we take the functions $\varphi_1(x)=\frac{2x^r}{1+x^r}$ and $\varphi_2(x)=x^s$ where $r>0,\ s>0$. Considering the fuzzy equivalence $E_{T_L}$, we obtain two fuzzy equivalencies by Proposition 2

$$E_1(x,y)=1-\left|\frac{2x^r}{1+x^r}-\frac{2y^r}{1+y^r}\right|, E_2(x,y)=1-\left|x^s-y^s\right|.$$

**Proposition 3.** *Let $\varphi$ be an automorphism of the unit interval and $E$ be a fuzzy equivalence. The function $E_\varphi:[0,1]^2\to[0,1]$ defined for all $x,y\in[0,1]$ by*

$$E_\varphi(x,y)=\varphi(E(x,y))$$

*is a fuzzy equivalence.*

**Example 5.** Considering the fuzzy equivalence $E_{T_P}$ and the automorphisms of the unit interval $\varphi_3(x) = \frac{2x}{1+x}, \varphi_4(x) = \frac{2x}{1+x^2}, \varphi_5(x) = \frac{2x^2}{1+x^2} and \varphi_6(x) = \frac{x}{2-x}$, we obtain four fuzzy equivalencies by Proposition 3

$$E_3(x,y) = \frac{2\min(x,y)}{x+y}, \qquad E_4(x,y) = \frac{2xy}{x^2+y^2},$$
$$E_5(x,y) = \frac{2\min(x^2,y^2)}{x^2+y^2}, \qquad E_6(x,y) = \frac{1-|x-y|}{1+|x-y|}.$$

**Proposition 4.** *Let $n$ be a fuzzy negation and $E$ be a fuzzy equivalence. The function $E_n : [0,1]^2 \to [0,1]$ defined for all $x, y \in [0,1]$ by*

$$E_n(x,y) = E(n(x), n(y))$$

*is a fuzzy equivalence.*

**Example 6.** Taking the fuzzy negation $n(x) = 1-x$ and the fuzzy equivalencies $E_{T_P}$ and $E_3$, we obtain two fuzzy equivalencies by Proposition 4

$$E_7(x,y) = \frac{\min(1-x,1-y)}{\max(1-x,1-y)}, \qquad E_8(x,y) = \frac{2\min(1-x,1-y)}{2-x-y}.$$

**Lemma 2.** *Let $T$ and $S$ be a t-norm and a t-conorm, respectively. Suppose $x_i, y_i, z_i \in [0,1]$ satisfying $\min(x_i, y_i) \geq z_i (i = 1,2,\ldots,n)$, then we have*

(1) $\min(T(x_1,x_2), T(y_1,y_2)) \geq T(z_1,z_2)$,

(2) $\min(S(x_1,x_2), S(y_1,y_2)) \geq S(z_1,z_2)$,

(3) $\min(\inf_{i\in I} x_i, \inf_{i\in I} y_i) \geq \inf_{i\in I} z_i$,

(4) $\min(\sup_{i\in I} x_i, \sup_{i\in I} y_i) \geq \sup_{i\in I} z_i$.

**Proposition 5.** *Let $E_p$ and $E_q$ be two fuzzy equivalencies and $T$ be a t-norm. The function $ET : [0,1]^2 \to [0,1]$ defined for all $x, y \in [0,1]$ by*

$$ET(x,y) = T(E_p(x,y), E_q(x,y))$$

*is a fuzzy equivalence.*

*Proof.* It is easy to see that ET satisfies E1, E2 and E3. We only prove E4 holds. Since $x \leq y \leq z$ implies that $\min(E_p(x,y), E_p(y,z)) \geq E_p(x,z)$, $\min(E_q(x,y), E_q(y,z)) \geq E_q(x,z)$, by Lemma 2 we have $\min(T(E_p(x,y), E_q(x,y)), T(E_p(y,z), E_q(y,z))) \geq T(E_p(x,z), E_q(x,z))$. Thus we obtain that $\min(ET(x,y), ET(y,z)) \geq ET(x,z)$ in case $x \leq y \leq z$. □

**Example 7.** Considering the fuzzy equivalencies $E_{T_P}$ and $E_7$ and the $t$-norm $T_P$, the corresponding fuzzy equivalence defined as Proposition 5 can be expressed as

$$E_9(x,y) = \frac{\min(x(1-y), y(1-x))}{\max(x(1-y), y(1-x))}.$$

Similar to Proposition 5 we obtain the following three propositions.

**Proposition 6.** *Let $E_p$ and $E_q$ be two fuzzy equivalencies and $S$ be a t-conorm. The function $ES : [0,1]^2 \to [0,1]$ defined for all $x, y \in [0,1]$by*

$$ES(x,y) = S(E_p(x,y), E_q(x,y))$$

*is a fuzzy equivalence.*

**Proposition 7.** *Let $(E_i)_{i \in I}$ be a family of fuzzy equivalencies. The function $E_I$ : $[0,1]^2 \to [0,1]$ defined for all $x, y \in [0,1]$ by*

$$E_I(x,y) = \inf_{i \in I} E_i(x,y)$$

*is a fuzzy equivalence.*

**Proposition 8.** *Let $(E_i)_{i \in I}$ be a family of fuzzy equivalencies. The function $E_S$ : $[0,1]^2 \to [0,1]$ defined for all $x, y \in [0,1]$ by*

$$E_S(x,y) = \sup_{i \in I} E_i(x,y)$$

*is a fuzzy equivalence.*

The following propositions show that some fuzzy equivalencies can be constructed by the same formula.

**Proposition 9.** *Suppose $E_\alpha$ is a function defined for all $x, y \in [0,1]$ by*

$$E_\alpha(x,y) = \frac{a - a\,|x-y| + b\min(x,y)}{a - (a-1)\,|x-y| + b\min(x,y)} \tag{3.1}$$

*with $a \geq 0, b \geq 0$ and $\max(a,b) \neq 0$, then $E_\alpha$ is a fuzzy equivalence.*

*Proof.* It is easy to prove that $E_\alpha$ is a function from $[0,1]^2$ to $[0,1]$ satisfying E1, E2 and E3. We only prove E4 holds. For all $x, y, z \in [0,1]$, if $x \geq y \geq z$, then we have

$$E_\alpha(x,y) = \frac{a - a(y-x) + bx}{a - (a-1)(y-x) + bx},$$

$$E_\alpha(y,z) = \frac{a - a(z-y) + by}{a - (a-1)(z-y) + by},$$

$$E_\alpha(x,z) = \frac{a - a(z-x) + bx}{a - (a-1)(z-x) + bx}.$$

Considering the following functions

$$f(t) = \frac{a - a(t-x) + bx}{a - (a-1)(t-x) + bx},$$

$$g(m) = \frac{a - a(z-m) + bm}{a - (a-1)(z-m) + bm},$$

where $t, m \in [0,1]$ and $t \geq x$, $m \leq z$, then we have

$$f'(t) = \frac{-a - bx}{[a - (a-1)(t-x) + bx]^2} \leq 0,$$

$$g'(m) = \frac{a + bz}{[a - (a-1)(z-m) + bm]^2} \geq 0.$$

Therefore, we can conclude that $f$ is decreasing with respect to $t$ and $g$ is increasing with respect to $m$. Thus we have $E_\alpha(x,y) \geq E_\alpha(x,z)$ and $E_\alpha(y,z) \geq E_\alpha(x,z)$ whenever $x \leq y \leq z$. □

*Remark* 2. In Eq.(3.1), if $a = \frac{1}{2}$, $b = 0$, then we have

$$E_\alpha(x,y) = \frac{1 - |x-y|}{1 + |x-y|} = E_6(x,y).$$

If $a = 1$, $b = 0$, then we have

$$E_\alpha(x,y) = 1 - |x-y| = E_{T_L}(x,y).$$

If $a = 0$, $b = 1$, then we have

$$E_\alpha(x,y) = \frac{\min(x,y)}{\max(x,y)} = E_{T_P}(x,y).$$

If $a = 0$, $b = 2$, then we have

$$E_\alpha(x,y) = \frac{2\min(x,y)}{x+y} = E_3(x,y).$$

**Proposition 10.** *Suppose $E_\beta$ is a function defined for all $x, y \in [0,1]$ by*

$$E_\beta(x,y) = \frac{a - a\,|x-y| + b\min(1-x, 1-y)}{a - (a-1)\,|x-y| + b\min(1-x, 1-y)} \tag{3.2}$$

*with $a \geq 0$, $b \geq 0$ and $\max(a,b) \neq 0$, then $E_\beta$ is a fuzzy equivalence.*

*Proof.* It is shown that $E_\beta(x,y) = E_\beta(1-x, 1-y)$ for all $x, y \in [0,1]$. Since $E_\beta$ is a fuzzy equivalence, by Proposition 4 we conclude that $E_\beta$ is a fuzzy equivalence. □

*Remark* 3. In Eq.(3.2), if $a = 0$, $b = 1$, then we have

$$E_\beta(x,y) = \frac{\min(1-x, 1-y)}{\max(1-x, 1-y)} = E_7(x,y).$$

If $a = 0$, $b = 2$, then we have

$$E_\beta(x,y) = \frac{2\min(1-x, 1-y)}{2 - x - y} = E_8(x,y).$$

**Lemma 3.** *Given a fuzzy equivalence $E$, for all $x, y, z \in [0,1]$ we have*

(1) $E(\max(x,z),\max(y,z)) \geq E(x,y)$,

(2) $E(\min(x,z),\min(y,z)) \geq E(x,y)$.

*Proof.* It is shown that there exist six permutations for $x$, $y$ and $z$, that is, $x \geq y \geq z$, $x \geq z \geq y$, $z \geq x \geq y$, $z \geq y \geq x$, $y \geq z \geq x$, $y \geq x \geq z$. It is enough to consider the preceding three cases, because $E$ is commutative. If $x \geq y \geq z$, then $E(\max(x,z),\max(y,z)) = E(x,y)$, $E(\min(x,z),\min(y,z)) = E(z,z) = 1 \geq E(x,y)$. If $x \geq z \geq y$, then $E(\max(x,z),\max(y,z)) = E(x,z) \geq E(x,y)$, $E(\min(x,z),\min(y,z)) = E(z,y) \geq E(x,y)$. If $z \geq x \geq y$, then $E(\max(x,z),\max(y,z)) = 1 \geq E(x,y)$, $E(\min(x,z),\min(y,z)) = E(x,y)$. □

## 3.2 Generating Dissimilarity Functions

In [8], Bustince et al. presented the definition of restricted dissimilarity function which arises from the concepts of dissimilarity and fuzzy equivalence. Then they constructed divergence measures by aggregating restricted dissimilarity functions.

**Definition 10.** A function $d : [0,1]^2 \to [0,1]$ is called a *restricted dissimilarity function* if it satisfies the following properties:

(1) $d(x,y) = d(y,x)$ for all $x, y \in [0,1]$,

(2) $d(x,y) = 0$ if and only if $x = y$ for all $x, y \in [0,1]$,

(3) $d(x,y) = 1$ if and only if $x = 0$ and $y = 1$ or $x = 1$ and $y = 0$,

(4) for all $x, y, z \in [0,1]$, if $x \leq y \leq z$, then $\max(d(x,y), d(y,z)) \leq d(x,z)$.

In the following, we propose a more general concept called dissimilarity function [28].

**Definition 11.** A function $d : [0,1]^2 \to [0,1]$ is called a *dissimilarity function* if it satisfies the following properties:

(d1) $d(x,y) = d(y,x)$ for all $x, y \in [0,1]$,

(d2) $d(x,x) = 0$ for all $x \in [0,1]$,

(d3) $d(1,0) = 1$,

(d4) for all $x, y, z \in [0,1]$, if $x \leq y \leq z$, then $\max(d(x,y), d(y,z)) \leq d(x,z)$.

*Remark* 4. d4 expresses the transitivity of dissimilarity functions. It says that the divergence between $x$ and $z$ exceeds that between $x$ and $y$ and that between $y$ and $z$ if the order relation $x \leq y \leq z$ holds. Let $T$ be a $t$-norm and $E_T$ be the biresiduation of $T$. Suppose $S$ is the $t$-conorm which is dual to $T$. If we define $d_T$ as $d_T(x,y) = 1 - E_T(x,y)$ for all $x, y \in [0,1]$, then it is proved that $d_T$ is a dissimilarity function. According to the $T$-transitivity of $E_T$, we have $S(d_T(x,y), d_T(y,z)) \geq d_T(x,z)$ for all $x, y, z \in [0,1]$. Hence, d4 is reduced to the inequality: for all $x \leq y \leq z$, $\max(d_T(x,y), d_T(y,z)) \leq d_T(x,z) \leq S(d_T(x,y), d_T(y,z))$.

It is shown that every restricted dissimilarity function in the sense of Bustince [8] is a dissimilarity function (the reciprocal is not true). Now we consider four further axioms in terms of dissimilarity functions.

(d5) $d(x,1) = 1 - x$ for all $x \in [0,1]$,

(d6) $d(x,0) = x$ for all $x \in [0,1]$,

(d7) $d(x,y) = 0$ if and only if x = y for all $x, y \in [0,1]$,

(d8) $d(x,y) = d(1-x, 1-y)$ for all $x, y \in [0,1]$.

**Proposition 11.** *Let $n$ be a fuzzy negation and $E$ be a fuzzy equivalence. Suppose $d_{n_1} : [0,1]^2 \to [0,1]$ is a function defined for all $x, y \in [0,1]$ by*

$$d_{n_1}(x,y) = n(E(x,y)).$$

*Then $d_{n_1}$ is a dissimilarity function.*

**Proposition 12.** *Let $n$ be a fuzzy negation and $d$ be a dissimilarity function. Suppose $d_{n_2} : [0,1]^2 \to [0,1]$ is a function defined for all $x, y \in [0,1]$ by*

$$d_{n_2}(x,y) = d(n(x), n(y)).$$

*Then $d_{n_2}$ is a dissimilarity function.*

*Remark* 5. In Propositions 11 and 12, if $n$, $E$ and $d$ are respectively strong fuzzy negations, restricted equivalence functions and restricted dissimilarity functions, then the obtained dissimilarity functions are restricted dissimilarity functions.

**Proposition 13.** *Let $\varphi$ be an automorphism of the unit interval and $d$ be a dissimilarity function. Suppose $d_{\varphi_1} : [0,1]^2 \to [0,1]$ is a function defined for all $x, y \in [0,1]$ by*

$$d_{\varphi_1}(x,y) = d(\varphi(x), \varphi(y)).$$

*Then $d_{\varphi_1}$ is a dissimilarity function.*

**Proposition 14.** *Let $\varphi$ be an automorphism of the unit interval and let $d$ be a dissimilarity function. Suppose $d_{\varphi_2} : [0,1]^2 \to [0,1]$ is a function defined for all $x, y \in [0,1]$ by*

$$d_{\varphi_2}(x,y) = \varphi(d(x,y)).$$

*Then $d_{\varphi_2}$ is a dissimilarity function.*

**Proposition 15.** *Let $\varphi$ be an automorphism of the unit interval and let $d$ be a dissimilarity function. Suppose $d_{\varphi_3} : [0,1]^2 \to [0,1]$ is a function defined for all $x, y \in [0,1]$ by*

$$d_{\varphi_3}(x,y) = \varphi^{-1}(d(x,y)).$$

*Then $d_{\varphi_3}$ is a dissimilarity function.*

**Example 8.** Let us take $\varphi_1$ and $\varphi_2$ in Propositions 13 and 15, respectively. Let $d_1(x,y) = |x-y|$ for all $x, y \in [0,1]$. By Propositions 13 and 15, we have $d_{\varphi_3}(x,y) = {\varphi_2}^{-1}(|\varphi_1(x) - \varphi_1(y)|)$, which is a restricted dissimilarity function presented in Proposition 2 of [8].

**Proposition 16.** *Let $d$ be a dissimilarity function and $E$ be a fuzzy equivalence. Suppose $d' : [0,1]^2 \to [0,1]$ is a function defined for all $x, y \in [0,1]$ by*

$$d'(x,\ y) = \frac{d(x,\ y)}{d(x,\ y)\ +\ E(x,\ y)}.$$

*Then $d'$ is a dissimilarity function.*

*Proof.* It is easy to prove that $d'$ is a function from $[0,1]^2$ to $[0,1]$ satisfying d1, d2 and d3. We only prove that $d'$ satisfies d4. For all $x, y, z \in [0,1]$ satisfying$x \le y \le z$, if $d'(x,\ y) = 0$, then $d'(x,\ y) \le d'(x,\ z)$. If $d'(x,\ y) \neq 0$, then we have $d'(x,\ z) \neq 0$. Therefore, we have

$$\frac{1}{d'(x,\ y)} = 1 + \frac{E(x,\ y)}{d(x,\ y)}, \frac{1}{d'(x,\ z)} = 1 + \frac{E(x,\ z)}{d(x,\ z)}.$$

Since $x \le y \le z$ implies $E(x,y) \ge E(x,z)$ and $d(x,y) \le d(x,z)$, we have $1 + \frac{E(x,\ y)}{d(x,\ y)} \ge 1 + \frac{E(x,\ z)}{d(x,\ z)}$. Therefore, we obtain that $d'(x,\ y) \le d'(x,\ z)$. The case of $d'(y,\ z) \le d'(x,\ z)$ can be proved similarly. □

**Proposition 17.** *Given a dissimilarity function $d$, for all $x, y \in [0,1]$, if $\left|x - \frac{1}{2}\right| \ge \left|y - \frac{1}{2}\right|$, then we have $d(x, 1-x) \ge d(y, 1-y)$.*

*Proof.* If $\left|x - \frac{1}{2}\right| \ge \left|y - \frac{1}{2}\right|$, then we obtain four permutations for $x$, $y$, $1-x$, $1-y$, i.e., $x \ge y \ge \frac{1}{2} \ge 1-y \ge 1-x$, $x \ge 1-y \ge \frac{1}{2} \ge y \ge 1-x$, $1-x \ge y \ge \frac{1}{2} \ge 1-y \ge x$, $1-x \ge 1-y \ge \frac{1}{2} \ge y \ge x$. It is easy to see that $d(x, 1-x) \ge d(y, 1-y)$ for these four cases. □

**Proposition 18.** *Given a dissimilarity function $d$, for all $x, y, z \in [0,1]$ we have:*

(1) $d(\max(x,z), \max(y,z)) \le d(x,y)$,

(2) $d(\min(x,z), \min(y,z)) \le d(x,y)$.

*Proof.* Let $E_n(x,y) = 1 - d(x,y)$ for all $x, y \in [0,1]$, then $E_n$ is a fuzzy equivalence. Thus the proof is immediate by Lemma 3. □

The following example shows that Lemma 3 and Proposition 18 may not hold for arbitrary *t*-norm or *t*-conorm.

**Example 9.** Let $E = E_2$ and $T(x,y) = xy$ for all $x, y \in [0,1]$. Suppose $x_0 \neq y_0$ and $0 < z_0 < 1$, we have $T(x_0, z_0) \neq T(y_0, z_0)$. Therefore, we have $E(T(x_0, z_0), T(y_0, z_0)) = \min(T(x_0, z_0), T(y_0, z_0)) = \min(x_0 z_0, y_0 z_0) < \min(x_0, y_0) = E(x_0, y_0)$. Let $E = E_6$ and $S(x,y) = \min(x+y, 1)$ for all $x, y \in [0,1]$. Let $x_1 = 0.6$,

$y_1 = 0.2$, $z_1 = 0.1$, then we have $S(x_1, z_1) = 0.7$, $S(y_1, z_1) = 0.3$. Therefore, $E(S(x_1, z_1), S(y_1, z_1)) = E(0.7, 0.3) = 0.3$ and $E(x_1, y_1) = 0.4$. Thus $E(S(x_1, z_1), S(y_1, z_1)) < E(x_1, y_1)$. This proves that $E(T(x, z), T(y, z)) \geq E(x, y)$ and $E(S(x, z), S(y, z)) \geq E(x, y)$ may not hold for any other $t$-norm or $t$-conorm that differ from $\min$ and $\max$. If we take $d(x, y) = 1 - E(x, y)$ for all $x, y \in [0, 1]$, then we can prove that $d(T(x, z), T(y, z)) \leq d(x, y)$ and $d(S(x, z), S(y, z)) \leq d(x, y)$ may not hold for any other $t$-norm or $t$-conorm that differ from $\min$ and $\max$.

**Proposition 19.** *Given a dissimilarity function $d$, for all $x, x^{'}, y, y^{'} \in [0, 1]$ we have:*

(1) $d(\max(x, y), \max(x^{'}, y^{'})) \leq \max(d(x, x^{'}), d(y, y^{'}))$,

(2) $d(\min(x, y), \min(x^{'}, y^{'})) \leq \min(d(x, x^{'}), d(y, y^{'}))$.

*Proof.* Let us denote $d(\max(x, y), \max(x^{'}, y^{'}))$ and $d(\min(x, y), \min(x^{'}, y^{'}))$ by $\Delta_1$ and $\Delta_2$, respectively. If $x \geq y$ and $x^{'} \geq y^{'}$, then $\Delta_1 = d(x, x^{'})$, $\Delta_2 = d(y, y^{'})$. If $x \leq y$ and $x^{'} \leq y^{'}$, then $\Delta_1 = d(y, y^{'})$, $\Delta_2 = d(x, x^{'})$. If $x \geq y$ and $x^{'} \leq y^{'}$, then six subcases should be considered, i.e., $x \geq y \geq y^{'} \geq x^{'}$, $x \geq y^{'} \geq y \geq x^{'}$, $x \geq y^{'} \geq x^{'} \geq y$, $y^{'} \geq x \geq x^{'} \geq y$, $y^{'} \geq x^{'} \geq x \geq y$, $y^{'} \geq x \geq y \geq x^{'}$. It is shown that $\Delta_1 = d(x, y^{'}) \leq \max(d(x, x^{'}), d(y, y^{'}))$ and $\Delta_2 = d(y, x^{'}) \leq \max(d(x, x^{'}), d(y, y^{'}))$ in these six subcases. The case of $x \leq y$ and $x^{'} \geq y^{'}$ can be proved similarly. □

## 3.3 Fuzzy Measures for the Comparison of Fuzzy Sets: Similarity Measures and Divergence Measures

In many fields of artificial intelligence, we need to compare the descriptions of objects. This comparison is frequently achieved through a measure intended to determine to which extent the descriptions have common points or differ from each other. In some occasions the comparison of fuzzy sets is done by quantifying the degree of equality or similarity between them, but in other cases we need to compare fuzzy sets by quantifying the degree of inequality or difference between them. Therefore, similarity measure and divergence measure are two basic fuzzy measures for the comparison of fuzzy sets.

The concept of similarity plays an essential role in human cognition because we often encounter situations where we have to distinguish between similar groups in day-to-day life. In fuzzy neural network, it is a common case to rule matching. This approach employs a similarity measure to determine whether a rule should be fired for a specific observation.

For example, in fuzzy pattern recognition, let $A_1, A_2, \ldots, A_n$ be $n$ fuzzy sets on the universal set $X$ for $n$ standard classes, $F(X)$ the class of all fuzzy sets of $X$, $A^c$ the complement of $A \in F(X)$. Given $A \in F(X)$, we need to know which class $A$ should be identified with. To solve this problem, we need a similarity measure to measure how close two fuzzy sets are. The measures used in order to make the comparison are normally demanded to fulfill the following natural properties.

(a) It is a non-negative and symmetric function of the two fuzzy sets to be compared.

(b) It becomes one when the two sets coincide.

(c) It becomes zero when the two sets are $X$ and the empty set, respectively.

(d) It increases when the two sets become "more similar" in some sense.

Several similarity measures for practical use have been proposed in the literature [4, 5, 10, 16, 20, 25, 27, 29, 30, 34, 35, 37, 38, 43]. When we review to these similarity measures, we find that most of them satisfy the above-mention conditions demanded from similarity measures. Therefore, we can select them to compare the similarity of two fuzzy sets in pattern recognition. Suppose $A$ and $B$ are two fuzzy sets, when measuring their similarity degree, different similarity measures will obtain different results, which result is the best choice of reflecting the intuitive closeness of those two fuzzy sets should be considered. Therefore, it is necessary to find more new similarity measures and give a deep comparison with other proposals in the literature.

Note that the similarity measures proposed and used in the literature are classified into three categories: (1) Distance based measures, (2) Set-theoretic based measures and (3) Implications based measures. However, there does not exist a unifying construction method for similarity measures. Section 3.3.1 will give two manners of generating similarity measures of Wang [38] from fuzzy equivalencies. One is by aggregating fuzzy equivalencies. The other is by reconstructing two formulae of fuzzy equivalencies. We will see from some examples that the similarity measures based on fuzzy equivalencies will serve as a very powerful tool for the calculation of similarity between fuzzy sets. They are general in the sense that by using different fuzzy equivalencies one gets the distance and the set-theoretic, as well as the implications based similarity measures designed in the literature.

In fuzzy set theory, several ways of measuring the difference between fuzzy sets by means of some functions have been proposed. It should be mentioned here that in some references these functions are called *distance measures* [3, 8, 15, 26, 28, 43], in some references these functions are called *dissimilarity measures* [4, 12], but in other references these functions are called *divergence measures* [1, 11, 32, 33]. The first axiomatic definition for distance measure was proposed by Liu in [43], where four axioms for distance measure were given. On the basis of exponential operation, Fan and Xie [15] defined a new distance measure. Bloch [3] classified distance measures into four categories through their construction approaches. Almost all the distance measures proposed in the literature can be classified into these four categories. In [4], Bouchon-Meunier et al. defined a $M$-measure of dissimilarity between fuzzy sets. This dissimilarity measure is indeed a distance measure in the sense of Liu [43] in case it is symmetrical and satisfies the triangular inequality. Based on logarithm operation, Bhandari and Pal [2] defined a divergence measure between fuzzy sets. In [32], Montes et al. introduced a new axiomatic definition for divergence measure. Their approach is based on three axioms modeling the minimal requirements for a function that tries to measure the separation or difference between fuzzy sets. In this chapter, we will call all functions measuring the difference between fuzzy sets divergence measures.

Based on the above-mentioned definitions of divergence measure, it is possible to define measures of divergence between fuzzy sets. When measuring the divergence degree between fuzzy sets, different divergence measures will obtain different results, which result is the best choice of reflecting the intuitive difference between fuzzy sets should be considered. Thus it is necessary to find some ways to construct more divergence measures and then compare them with other proposals in the literature. Montes et al. [32, 33] proposed a way of constructing divergence measure. This kind of divergence measure is also a divergence measure in the sense of Liu [43], based on fuzzy union and intersection. Bustince et al. [8] normalized the definition of divergence measure in the sense of Liu and presented a method to construct normal divergence measures. Both of Montes' and Bustince's approaches to constructing divergence measures are based on the use of restricted dissimilarity functions. Although Montes' and Bustince's approaches can be used to construct numerous divergence measures, including many existing ones, there still exist some kinds of divergence measures that cannot be constructed by their approaches (e.g., divergence measures based on the set theoretic approach [3]). Section 3.3.2 will propose two approaches to constructing divergence measure in the sense of Liu and that in the sense of Montes. The construction is based on the use of dissimilarity functions and fuzzy equivalencies.

### 3.3.1 Constructing Similarity Measures via Fuzzy Equivalencies

The concept of similarity measure has been quite studied in the fuzzy literature, like for example those of Pappis and Karacapilidis [35], or that of Wang et al. [40], or that of Liu [43]; however, the authors who work on this theme do not agree on the axioms that must be demanded from such functions. According to the four natural properties for similarity measures, we note that a reasonable similarity measure used for pattern recognition must satisfy at least the following three conditions [38].

**Definition 12.** A function $N : F(X) \times F(X) \to [0,1]$ is called a *similarity measure* if it satisfies the following properties:

(N1) $N(X, \varnothing) = 0$ and $N(A, A)$ =1 whenever $A \in F(X)$,

(N2) $N(A, B) = N(B, A)$ whenever $A, B \in F(X)$,

(N3) for all $A, B, C \in F(X), N(A, C) \leq \min(N(A, B), N(B, C))$ whenever $A \subseteq B \subseteq C$.

This subsection will introduce two manners of constructing similarity measures.

**Proposition 20.** *Given a discrete universe $X = \{x_1, x_2, ..., x_n\}$. Let $E$ be a fuzzy equivalence, $M$ be an aggregation operator. Suppose $N_\tau : F(X) \times F(X) \to [0,1]$ is a function defined for all $A, B \in F(X)$ by*

$$N_\tau(A, B) = \mathop{\mathrm{M}}_{i=1}^{n} E(A(x_i), B(x_i)),$$

*then $N_\tau$ is a similarity measure.*

*Proof.* It is easy to see that the defined function $N_\tau$ satisfies N1 and N2. N3: Since $A \subseteq B \subseteq C$ implies$A(x_i) \leq B(x_i) \leq C(x_i)$ for each $x_i \in X$, by E4 we have

$\min(E(A(x_i), B(x_i)), E(B(xi), C(x_i))) \geq E(A(x_i), C(x_i))$. According to M4 we have

$$\min(\mathop{\mathrm{M}}_{i=1}^{n} E(A(x_i), B(x_i)), \mathop{\mathrm{M}}_{i=1}^{n} E(B(x_i), C(x_i))) \geq \mathop{\mathrm{M}}_{i=1}^{n} E(A(x_i), C(x_i)).$$

Therefore, we have $\min(N_\tau(A, B), N_\tau(B, C)) \geq N_\tau(A, C)$. □

*Remark* 6. In Proposition 20, if we take the arithmetic mean aggregation operator $M$, then we have that for all $A, B \in F(X)$

$$N_\tau(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} E(A(x_i), B(x_i)). \tag{3.3}$$

Compared with other aggregation operators, the arithmetic mean one has the significance of average. In Eq.(3.3), each element for $A$ and $B$ plays an equal role in the similarity comparison. Thus we will discuss the similarity measures defined by Eq.(3.3) in this chapter.

**Example 10.** Consider the fuzzy equivalencies $E_{T_L}, E_{T_P}$ and E1-E9 with $r = 1$ and $s = 2$ in $E_1$ and $E_2$, respectively. The corresponding similarity measures defined in Eq.(3.3) can be expressed as

$$N_1(A, B) = 1 - \frac{1}{n} \sum\nolimits_{i=1}^{n} |A(x_i) - B(x_i)|,$$

$$N_2(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{\min(A(x_i), B(x_i))}{\max(A(x_i), B(x_i))},$$

$$N_3(A, B) = 1 - \frac{1}{n} \sum\nolimits_{i=1}^{n} \left| \frac{2A(x_i)}{1 + A(x_i)} - \frac{2B(x_i)}{1 + B(x_i)} \right|,$$

$$N_4(A, B) = 1 - \frac{1}{n} \sum\nolimits_{i=1}^{n} \left| A(x_i)^2 - B(x_i)^2 \right|,$$

$$N_5(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{2\min(A(x_i), B(x_i))}{A(x_i) + B(x_i)},$$

$$N_6(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{2A(x_i)B(x_i)}{A(x_i)^2 + B(x_i)^2},$$

$$N_7(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{2\min(A(x_i)^2, B(x_i)^2)}{A(x_i)^2 + B(x_i)^2},$$

$$N_8(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{1 - |A(x_i) - B(x_i)|}{1 + |A(x_i) - B(x_i)|},$$

$$N_9(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{\min(1 - A(x_i), 1 - B(x_i))}{\max(1 - A(x_i), 1 - B(x_i))},$$

$$N_{10}(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{2\min(1 - A(x_i), 1 - B(x_i))}{2 - A(x_i) - B(x_i)},$$

$$N_{11}(A, B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{\min((1 - A(x_i))B(x_i), (1 - B(x_i))A(x_i))}{\max((1 - A(x_i))B(x_i), (1 - B(x_i))A(x_i))}.$$

*Remark* 7. Note that $N_1$ was first given in [39], $N_2$, $N_5$ and $N_9$ were first given in [14], $N_8$ was first given in [19]. These five similarity measures were proposed through different approaches in the literature. However, all of them can be constructed by the same formula in this chapter.

We have proved that $E_\alpha$ and $E_\beta$ are fuzzy equivalencies in Propositions 9 and 10. Therefore, we can use them to construct similarity measures by Eq.(3.3). The following two propositions show that we can obtain another kind of general formula of similarity measure by reconstructing $E_\alpha$ and $E_\beta$.

**Proposition 21.** *Given a discrete universe $X = \{x_1, x_2, ..., x_n\}$. Suppose that $N_\theta$ is a function defined for all $A, B \in F(X)$ by*

$$N_\theta(A,B) = \frac{\sum_{i=1}^{n}(a - a\,|A(x_i) - B(x_i)| + b\min(A(x_i), B(x_i)))}{\sum_{i=1}^{n}(a - (a-1)\,|A(x_i) - B(x_i)| + b\min(A(x_i), B(x_i)))}. \qquad (3.4)$$

*with $a \geq 0$, $b \geq 0$ and $\max(a,b) \neq 0$, then $N_\theta$is a similarity measure.*

*Proof.* It is easy to see that $N_\theta$ is a function from $F(X) \times F(X)$ to $[0,1]$ satisfying N1 and N2. We only prove $N_\theta$ satisfies N3 here.

Since $A \subseteq B \subseteq C$ implies $A(x_i) \leq B(x_i) \leq C(x_i)$ for each $x_i \in X$, we have

$$N_\theta(A,B) = \frac{\sum_{i=1}^{n}(a - a(B(x_i) - A(x_i)) + bA(x_i))}{\sum_{i=1}^{n}(a - (a-1)(B(x_i) - A(x_i)) + bA(x_i))},$$

$$N_\theta(B,C) = \frac{\sum_{i=1}^{n}(a - a(C(x_i) - B(x_i)) + bB(x_i))}{\sum_{i=1}^{n}(a - (a-1)(C(x_i) - B(x_i)) + bB(x_i))},$$

$$N_\theta(A,C) = \frac{\sum_{i=1}^{n}(a - a(C(x_i) - A(x_i)) + bA(x_i))}{\sum_{i=1}^{n}(a - (a-1)(C(x_i) - A(x_i)) + bA(x_i))}.$$

Considering the following two functions

$$f(y_1, y_2, ..., y_n) = \frac{\sum_{i=1}^{n}(a - a(y_i - A(x_i)) + bA(x_i))}{\sum_{i=1}^{n}(a - (a-1)(y_i - A(x_i)) + bA(x_i))},$$

$$g(w_1, w_2, ..., w_n) = \frac{\sum_{i=1}^{n}(a - a(C(x_i) - w_i) + bw_i)}{\sum_{i=1}^{n}(a - (a-1)(C(x_i) - w_i) + bw_i)},$$

where $y_i, w_i \in [0,1]$ and $y_i \geq A(xi)$, $w_i \leq C(x_i)$ for each $x_i \in X$, then we have

$$\begin{aligned}
\frac{\partial f}{\partial y_i} &= \frac{-\sum_{i=1}^{n}(a + bA(x_i))}{\left[\sum_{i=1}^{n}(a - (a-1)(y_i - A(x_i)) + bA(x_i))\right]^2} \leq 0, \\
\frac{\partial g}{\partial w_i} &= \frac{\sum_{i=1}^{n}(a + bC(x_i))}{\left[\sum_{i=1}^{n}(a - (a-1)(C(x_i) - w_i) + bw_i)\right]^2} \geq 0.
\end{aligned}$$

Therefore, we can conclude that $f$ is decreasing with respect to $y_i (i = 1, 2, \ldots, n)$ and $g$ is increasing with respect to $w_i (i = 1, 2, \ldots, n)$. Thus we have $N_\theta(A,B) \geq N_\theta(A,C)$ and $N_\theta(B,C) \geq N_\theta(A,C)$ whenever $A \subseteq B \subseteq C$. $\square$

*Remark* 8. In Eq.(3.4), let $a = \frac{1}{2}, b = 0$, then we have

$$N_{12}(A,B) = \frac{\sum_{i=1}^{n}(1 - |A(x_i) - B(x_i)|)}{\sum_{i=1}^{n}(1 + |A(x_i) - B(x_i)|)}.$$

Let $a = 1$, $b = 0$, then we have

$$N_{13}(A,B) = 1 - \frac{1}{n}\sum_{i=1}^{n}|A(x_i) - B(x_i)| = N_1(A,B).$$

Let $a = 0$, $b = 1$, then we have

$$N_{14}(A,B) = \frac{\sum_{i=1}^{n}\min(A(x_i), B(x_i))}{\sum_{i=1}^{n}\max(A(x_i), B(x_i))}.$$

Let $a = 0$, $b = 2$, then we have

$$N_{15}(A,B) = \frac{2\sum_{i=1}^{n}\min(A(x_i), B(x_i))}{\sum_{i=1}^{n}(A(x_i) + B(x_i))}.$$

Let $a = 1$, $b = 2$, then we have

$$N_{16}(A,B) = \frac{\sum_{i=1}^{n}(1 - |A(x_i) - B(x_i)| + 2\min(A(x_i), B(x_i)))}{\sum_{i=1}^{n}(1 + 2\min(A(x_i), B(x_i)))}.$$

Note that the similarity measure $N_{12}$ was designed in [19] and $N_{14}$, $N_{15}$ were designed in [35], where the computation was more complicated than the one given before.

Similar to Proposition 21 we obtain the following proposition.

**Proposition 22.** *Given a discrete universe $X = \{x_1, x_2, ..., x_n\}$. Suppose that $N_\delta$ is a function defined for all $A, B \in F(X)$ by*

$$N_\delta(A,B) = \frac{\sum_{i=1}^{n}(a - a\,|A(x_i) - B(x_i)| + b\min(1 - A(x_i), 1 - B(x_i)))}{\sum_{i=1}^{n}(a - (a-1)\,|A(x_i) - B(x_i)| + b\min(1 - A(x_i), 1 - B(x_i)))}, \tag{3.5}$$

*with $a \geq 0$, $b \geq 0$ and $\max(a,b) \neq 0$, then $N_\delta$ is a similarity measure.*

*Remark* 9. In Eq.(3.5), let $a = 0$, $b = 1$, then we have

$$N_{17}(A,B) = \frac{\sum_{i=1}^{n}\min(1 - A(x_i), 1 - B(x_i))}{\sum_{i=1}^{n}\max(1 - A(x_i), 1 - B(x_i))}.$$

Let $a = 0$, $b = 2$, then we have

$$N_{18}(A,B) = \frac{2\sum_{i=1}^{n}\min(1 - A(x_i), 1 - B(x_i))}{\sum_{i=1}^{n}(2 - A(x_i) - B(x_i))}.$$

Note that the similarity measure $N_{17}$ was designed in [16].

### 3.3.2 Constructing Divergence Measures via Dissimilarity Functions

Let us consider the definition of divergence measure in the sense of Liu [43] and that in the sense of Montes [32].

**Definition 13.** A function $D : F(X)\times F(X) \to [0, \infty)$ is called a *divergence measure* in the sense of Liu if it satisfies:

(1) $D(A, B) = D(B, A)$ for all $A, B \in F(X)$,

(2) $D(A, A) = 0$ for all $A \in F(X)$,

(3) for all $A, B, C \in F(X)$, if $A \subseteq B \subseteq C$, then $\max(D(A, B), D(B, C)) \leq D(A, C)$,

(4) $D(A, B) \leq D(P, P^c)$for all $A, B \in F(X)$ and $P \in P(X)$.

**Definition 14.** A function $D : F(X)\times F(X) \to [0, \infty)$ is called a *divergence measure* in the sense of Montes if it satisfies:

(1) $D(A, B) = D(B, A)$ for all $A, B \in F(X)$,

(2) $D(A, A) = 0$ for all $A \in F(X)$,

(3) $\max(D(A \cup C, B \cup C), D(A \cap C, B \cap C)) \leq D(A, B)$ for all $A, B, C \in F(X)$.

The first three axioms in both of the definitions are defined on the basis of the following natural properties [32]:

(a) It is a non-negative and symmetric function of the compared fuzzy sets.

(b) It becomes zero when the compared fuzzy sets coincide.

(c) It decreases when the compared fuzzy sets become "more similar" in some sense.

The fourth axiom in Definition 13 is defined on the basis of the following natural property.

(d) For a crisp set $P \in P(X)$, we have $P(x) = 1$ or $0$ for any $x \in X$. Hence, $P^c(x) = 0$ or $1$. Thus for any $x \in X$, we have $x \in P$ or $x \in P^c$. Therefore, the normalized divergence degree between $P$ and $P^c$ is always equal to $1$. However, the normalized divergence degree between any two fuzzy sets $A$ and $B$ is not always equal to $1$, i.e., it is less than or equal to $1$. Therefore, the divergence between a crisp set and its complement is larger than the divergence between any two fuzzy sets.

The third axiom of the two definitions depends on the formalization of the concept of "more similar". Liu [43] based his approach on the fact that the closer two fuzzy sets are, the smaller their divergence is. Montes [32] based her approach on the fact that if joining (resp. intersecting) both $A$ and $B$ with another fuzzy set $C$, the divergence should decrease, as $A \cup C$ and $B \cup C$ (resp. $A \cap C$ and $B \cap C$) are more similar than $A$ and $B$. It is shown that Axiom 3 in the sense of Liu is more general than that in the sense of Montes from the following proposition [32].

**Proposition 23.** *Let $D$ be a divergence measure in the sense of Montes, then*

$$\max(D(A,B),D(B,C)) \le D(A,C),$$

*whenever $A \subseteq B \subseteq C$.*

This subsection will introduce two manners of constructing divergence measure in the sense of Liu and that in the sense of Montes.

**Proposition 24.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $d$ be a dissimilarity function. Suppose $D : F(X) \times F(X) \to [0,\infty)$ is a function defined for all $A, B \le F(X)$ by*

$$D(A,B) = \sum\nolimits_{i=1}^{n} d(A(x_i),B(x_i)).$$

*Then $D$ is both a divergence measure in the sense of Montes and a divergence measure in the sense of Liu.*

*Proof.* It is quite evident that $D$ satisfies Axioms 1 and 2. Note that

$$D(A\cup C,B\cup C) = \sum\nolimits_{i=1}^{n} d(\max(A(x_i),C(x_i)),\max(B(x_i),C(x_i))),$$

$$D(A\cap C,B\cap C) = \sum\nolimits_{i=1}^{n} d(\min(A(x_i),C(x_i)),\min(B(x_i),C(x_i))).$$

Since $d$ is a dissimilarity function, according to Proposition 19, we have

$$\sum\nolimits_{i=1}^{n} d(\max(A(x_i),C(x_i)),\max(B(x_i),C(x_i))) \le \sum\nolimits_{i=1}^{n} d(A(x_i),B(x_i)),$$

$$\sum\nolimits_{i=1}^{n} d(\min(A(x_i),C(x_i)),\min(B(x_i),C(x_i))) \le \sum\nolimits_{i=1}^{n} d(A(x_i),B(x_i)).$$

Therefore, we have $\max(D(A\cup C,B\cup C),D(A\cap C,B\cap C)) \le D(A,B)$. According to Proposition 23, we have $\max(D(A,B),D(B,C)) \le D(A,C)$ whenever $A \subseteq B \subseteq C$. Furthermore, we have $D(P,P^c) = \sum_{i=1}^{n} d(P(x_i),1-P(x_i)) = n \ge D(A,B)$ for all $A,B \in F(X)$. Thus we conclude that $D$ is both a divergence measure in the sense of Montes and a divergence measure in the sense of Liu. □

The result given in Proposition 24 is just the way of constructing divergence measures proposed by Montes [32] and Bustince [8].

**Proposition 25.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $d$ and $E$ be a dissimilarity function and a fuzzy equivalence, respectively. If $D$ is a function defined for all $A, B \in F(X)$ by*

$$D(A,B) = \frac{a\sum_{i=1}^{n} d(A(x_i),B(x_i))}{\begin{array}{r} a\sum_{i=1}^{n} d(A(x_i),B(x_i)) + b\sum_{i=1}^{n} E(A(x_i),B(x_i)) \\ + c\sum_{i=1}^{n} \min(A(x_i),B(x_i)) \end{array}}, \tag{3.6}$$

*where $a > 0$, $b \ge 0$, $c \ge 0$ and $\max(b,c) > 0$, then $D$ is a divergence measure in the sense of Liu.*

*Proof.* It is easy to prove that $D$ is a function from $F(X) \times F(X)$ to $[0, 1]$ satisfying Axioms 1, 2 and 4. We prove that Axiom 3 holds.

For all $A, B, C \in F(X)$ satisfying $A \subseteq B \subseteq C$, if $D(A, B) = 0$, then $D(A, B) \leq D(A, C)$. Since $A \subseteq B \subseteq C$ implies $A(x_i) \leq B(x_i) \leq C(x_i)$ for each $x_i \in X$, we have $d(A(x_i), B(x_i)) \leq d(A(x_i), C(x_i))$. Thus we obtain that $\sum_{i=1}^{n} d(A(x_i), B(x_i)) \leq \sum_{i=1}^{n} d(A(x_i), C(x_i))$.If $D(A, B) \neq 0$, then $D(A, C) \neq 0$. Therefore, we have

$$\frac{1}{D(A,B)} = 1 + \frac{b \sum_{i=1}^{n} E(A(x_i), B(x_i)) + c \sum_{i=1}^{n} A(x_i)}{a \sum_{i=1}^{n} d(A(x_i), B(x_i))},$$

$$\frac{1}{D(A,C)} = 1 + \frac{b \sum_{i=1}^{n} E(A(x_i), C(x_i)) + c \sum_{i=1}^{n} A(x_i)}{a \sum_{i=1}^{n} d(A(x_i), C(x_i))}.$$

Since $A(x_i) \leq B(x_i) \leq C(x_i)$ for each $x_i \in X$, we have

$$b \sum_{i=1}^{n} E(A(x_i), B(x_i)) \geq b \sum_{i=1}^{n} E(A(x_i), C(x_i)),$$

$$a \sum_{i=1}^{n} d(A(x_i), B(x_i)) \leq a \sum_{i=1}^{n} d(A(x_i), C(x_i)).$$

Thus we have $\frac{1}{D(A,B)} \geq \frac{1}{D(A,C)}$.Therefore, $D(A, B) \leq D(A, C)$ holds. The case of $D(B, C) \leq D(A, C)$ can be proved similarly. □

**Example 11.** Let us take $d = d_1$ and $E = E_1$ in Proposition 25.

(1) If $a = 1$, $b = 1$, $c = 0$, then we have

$$D_1(A, B) = \frac{1}{n} \sum_{i=1}^{n} |A(x_i) - B(x_i)|.$$

It is the normalization of Hamming distance.

(2) If $a = 1$, $b = 0$, $c = 1$, then we have

$$D_2(A, B) = \frac{\sum_{i=1}^{n} |A(x_i) - B(x_i)|}{\sum_{i=1}^{n} \max(A(x_i), B(x_i))}.$$

It is a divergence measure constructed by the set theoretic approach [3].

(3) If $a = 1$, $b = 0$, $c = 2$, then we have

$$D_3(A, B) = \frac{\sum_{i=1}^{n} |A(x_i) - B(x_i)|}{\sum_{i=1}^{n} (A(x_i) + B(x_i))}.$$

It is a divergence measure proposed by Pappis and Karacapilidis [35].

Note that Eq.(3.6) is a unified form of $D_1$, $D_2$ and $D_3$. We can discuss properties of them through discussing properties of Eq.(3.6). More divergence measures can be obtained in case we take different $a$, $b$, $c$ on Eq.(3.6). The following example shows that the proposed divergence measure is more powerful than the Hamming distance which has been extensively used.

**Example 12.** Let $X = \{x_1, x_2\}$ and let $A$, $B$ and $C$ be fuzzy sets in $F(X)$ with $A = \{0, 1\}$, $B = \left\{\frac{1}{2}, 1\right\}$, $C = \left\{0, \frac{1}{2}\right\}$. Using the Hamming distance, we have $D(A, B) = \sum_{i=1}^{2} |A(x_i) - B(x_i)| = \frac{1}{2} = D(A, C)$. The divergence between $A$ and $B$ is the same as the one between $A$ and $C$. On the other hand, the $\alpha$-cuts of $A$ and $B$ for $\alpha > \frac{1}{2}$(i.e., the "important" $\alpha$-cuts) are the same, which is not true for $A$ and $C$. From this point of view we can require that the divergence between $A$ and $B$ should be smaller than that between $A$ and $C$ (this viewpoint was given by Couso [11]). We try to introduce another divergence measure that would reflect the above-mentioned difference. Let us take $d = d_1$ and $E = E_1$ on Eq.(3.6) ($c \neq 0$), then we have

$$\begin{aligned} D(A, B) &= \frac{a \sum_{i=1}^{2} |A(x_i) - B(x_i)|}{\begin{array}{r} a \sum_{i=1}^{2} |A(x_i) - B(x_i)| + b \sum_{i=1}^{2} (1 - |A(x_i) - B(x_i)|) \\ + c \sum_{i=1}^{2} \min(A(x_i), B(x_i)) \end{array}} \\ &= \frac{a}{a + 3b + 2c}, \end{aligned}$$

$$\begin{aligned} D(A, C) &= \frac{a \sum_{i=1}^{2} |A(x_i) - C(x_i)|}{\begin{array}{r} a \sum_{i=1}^{2} |A(x_i) - C(x_i)| + b \sum_{i=1}^{2} (1 - |A(x_i) - C(x_i)|) \\ + c \sum_{i=1}^{2} \min(A(x_i), C(x_i)) \end{array}} \\ &= \frac{a}{a + 3b + c}. \end{aligned}$$

Therefore, we have $D(A, C) \geq D(A, B)$, what reflects the fact that the divergence between $A$ and $C$ are on higher membership degree.

*Remark* 10. By Eq.(3.6), we obtain divergence measures based on the set theoretic approach.

**Proposition 26.** *Let $A$, $B$ and $C$ be fuzzy sets in $F(X)$, $D$ a function defined by Eq.(3.6), then we have $D(A \cup C, B \cup C) \leq D(A, B)$.*

*Proof.* If $D(A \cup C, B \cup C) = 0$, then we have $D(A, B) \geq D(A \cup C, B \cup C)$. If $D(A \cup C, B \cup C) \neq 0$, then $D(A, B) \neq 0$. Therefore, we have

$$\frac{1}{D(A, B)} = 1 + \frac{b \sum_{i=1}^{n} E(A(x_i), B(x_i)) + c \sum_{i=1}^{n} \min(A(x_i), B(x_i))}{a \sum_{i=1}^{n} d(A(x_i), B(x_i))},$$

$$\frac{1}{D(A \cup C, B \cup C)} = 1 + \frac{b \sum_{i=1}^{n} E(M(x_i), Q(x_i)) + c \sum_{i=1}^{n} \min(M(x_i), Q(x_i))}{a \sum_{i=1}^{n} d(M(x_i), Q(x_i))},$$

where $M(x_i) = \max(A(x_i), C(x_i))$, $Q(x_i) = \max(B(x_i), C(x_i))$ for each $x_i \in X$. By Lemma 3 and Proposition 19, we have

$$b \sum_{i=1}^{n} E(M(x_i), Q(x_i)) \geq b \sum_{i=1}^{n} E(A(x_i), B(x_i)),$$

$$a \sum_{i=1}^{n} d(M(x_i), Q(x_i)) \leq a \sum_{i=1}^{n} d(A(x_i), B(x_i)).$$

Thus $D(A, B) \geq D(A \cup C, B \cup C)$. $\square$

*Remark* 11. For the function defined by Eq.(3.6), we cannot always conclude that $D(A\cap C, B\cap C)\le D(A,B)$. For example, suppose $X=\{x_1,x_2\}$, considering the divergence measure in the sense of Liu

$$D_2(A,B)=\frac{\sum_{i=1}^{n}|A(x_i)-B(x_i)|}{\sum_{i=1}^{n}\max(A(x_i),B(x_i))}.$$

Let $A=\{0.7,1\}$, $B=\{0.8,0.4\}$, $C=\{0.3,0.9\}$, then we have $D_2(A\cap C,B\cap C)=0.42$, $D_2(A,B)=0.39$ and thus $D_2(A,B)<D_2(A\cap C,B\cap C)$. Therefore, the function defined by Eq.(3.6) is not always a divergence measure in the sense of Montes.

The following proposition gives the conditions under which Eq.(3.6) satisfies $D(A\cap C,B\cap C)\le D(A,B)$.

**Proposition 27.** *Let $A$, $B$ and $C$ be fuzzy sets in $F(X)$, $D$ a function defined by Eq.(3.6). If $c=0$ in Eq.(3.6), then we have $D(A\cap C,B\cap C)\le D(A,B)$.*

*Proof.* It can be proved in the same manner with Proposition 26. □

According to Propositions 26 and 27, we have

**Proposition 28.** *Given a discrete universe $X=\{x_1,x_2,\ldots,x_n\}$. Let $d$ and $E$ be a dissimilarity function and a fuzzy equivalence, respectively. If $D$ is a function defined for all $A,B\in F(X)$ by*

$$D(A,B)=\frac{a\sum_{i=1}^{n}d(A(x_i),B(x_i))}{a\sum_{i=1}^{n}d(A(x_i),B(x_i))+b\sum_{i=1}^{n}E(A(x_i),B(x_i))}, \tag{3.7}$$

*where $a>0$, $b>0$, then $D$ is a divergence measure in the sense of Montes.*

*Remark* 12. We find from Propositions 25 and 28 that the divergence measure in the sense of Liu is more general than that in the sense of Montes.

**Example 13.** Let us take $d=d_1$ and $E=E_1$ in Proposition 28. If $a=1$, $b=\frac{1}{2}$, then we have

$$D_4(A,B)=\frac{2\sum_{i=1}^{n}|A(x_i)-B(x_i)|}{\sum_{i=1}^{n}(1+|A(x_i)-B(x_i)|)}.$$

It is a divergence measure in the sense of Montes as well as a divergence measure in the sense of Liu.

## 3.4 Relationships Between Similarity Measures and Subsethood Measures

This subsection discusses relationships between similarity measures and subsethood measures and proposes several propositions that similarity measures and subsethood measures can be transformed by each other based on their axiomatic definitions.

We have referred to the definition of similarity measure in Section 3.3. Several reasonable properties may be required for similarity measures; among them we consider the following in this subsection:

(N4) $N(A, B) = 1$ if and only if $A = B$,

(N5) $N(A, B) = 0$ if and only if $A \cap B = \emptyset$ and $A \cup B \neq \emptyset$,

(N6) if $A \subseteq B$, then $N(A \cup C, A) \leq N(B \cup C, B)$ for all $A, B, C \in F(X)$,

(N7) if $A \subseteq B$, then $N(A, A \cap C) \geq N(B, B \cap C)$ for all $A, B, C \in F(X)$,

(N8) $N(A, A \cap A^c) = 0$ if and only if $A = X$ for all $A \in F(X)$

(N9) $N(A \cup A^c, A^c) = 0$ if and only if $A = X$ for all $A \in F(X)$.

On the basis of Kosko's [23, 24] subsethood measure, fuzzy entropy and Wilmott's work [42], Young defined subsethood measure in the following way:

**Definition 15.** A function $c_{VY} : F(X) \times F(X) \to [0, 1]$ is called a *VY-subsethood measure*, if $c_{VY}$ satisfies the following conditions:

(C1) $c_{VY}(A, B) = 1$ if and only if $A \subseteq B$, i.e., $A(x) \leq B(x)$ for all $x \in X$,

(C2) if $[\frac{1}{2}] \subseteq A$, then $c_{VY}(A, A^c) = 0$ if and only if $A = X$,

(C3) if $A \subseteq B \subseteq C$, then $c_{VY}(C, A) \leq c_{VY}(B, A)$ and if $A \subseteq B$, then $c_{VY}(C, A) \leq c_{VY}(C, B)$.

It has been pointed out in [17] that C3 is too strong when considering the relation between subsethood measure and fuzzy entropy. Therefore, Fan et al. [17] thought a simpler form of definition of subsethood measure based on Young's definition.

**Definition 16.** A function $c_* : F(X) \times F(X) \to [0, 1]$ is called a $*$*-subsethood measure*, if $c_*$ satisfies the following conditions:

(C1) $c_*(A, B) = 1$ if and only if $A \subseteq B$, i.e., $A(x) \leq B(x)$ for all $x \in X$,

(C2) if $[\frac{1}{2}] \subseteq A$, then $c_*(A, A^c) = 0$ if and only if $A = X$,

(C3) if $A \subseteq B \subseteq C$, then $c_*(C, A) \leq c_*(B, A)$ and $c_*(C, A) \leq c_*(C, B)$.

Obviously, the only difference between Definitions 15 and 16 is in C3 where Young demands increasingness in the second component. Thus every *VY*-subsethood measure is also a $*$-subsethood measure. In 2006, Bustince et al. [9] modified two of Young's axioms and proposed a new class of subsethood measure called *DI-subsethood measure*.

**Definition 17.** A function $c_{DI} : F(X) \times F(X) \to [0, 1]$ is called a *DI-subsethood measure*, if $c_{DI}$ satisfies the following conditions:

(C1) $c_{DI}(A, B) = 1$ if and only if $A \subseteq B$, i.e., $A(x) \leq B(x)$ for all $x \in X$,

(C2) $c_{DI}(A, A^c) = 0$ if and only if $A = X$ ,

(C3) if $A \subseteq B$, then $c_{DI}(A, C) \geq c_{DI}(B, C)$ and $c_{DI}(C, A) \leq c_{DI}(C, B)$.

It is shown that every *DI*-subsethood measure is a *VY*-subsethood measure and therefore, it is also a $*$-subsethood measure.

Considering that an axiom definition must generally be abstract and simple, Fan [17] gave the following definition.

**Definition 18.** A function $c : F(X) \times F(X) \to [0,1]$ is called a *subsethood measure*, if c satisfies the following conditions:

(C1) $c(A,B) = 1$ in case $A \subseteq B$,

(C2) $c(X, \emptyset) = 0$,

(C3) if $A \subseteq B \subseteq C$, then $c(C,A) \leq c(B,A)$ and $c(C,A) \leq c(C,B)$.

We can see that *VY*-subsethood measure, $*$-subsethood measure and *DI*-subsethood measure are special cases of subsethood measures. In the following we derive subsethood measures from similarity measures.

**Proposition 29.** *Given a discrete universe* $X = \{x_1, x_2, \ldots, x_n\}$*. Let* $N$ *be a similarity measure,* $c$ *a function defined for all* $A, B \in F(X)$ *by* $c(A,B) = N(A, A \cap B)$*. Then we obtain the following conclusions:*

(1) $c$ is a subsethood measure,

(2) if $N$ satisfies N4 and N8, then $c$ is a *VY*-subsethood measure,

(3) if $N$ satisfies N4 and N8, then $c$ is a $*$-subsethood measure,

(4) if $N$ satisfies N4 and N5, then $c$ is a *VY*-subsethood measure,

(5) if $N$ satisfies N4 and N5, then $c$ is a $*$-subsethood measure,

(6) if $N$ satisfies N4, N7 and N8, then $c$ is a *DI*-subsethood measure.

*Proof.*
(1) **(C1)** If $A \subseteq B$, then $c(A,B) = N(A, A \cap B) = N(A,A) = 1$. **(C2)** $c(X, \emptyset) = N(X, X \cap \emptyset) = N(X, \emptyset) = 0$. **(C3)** If $A \subseteq B \subseteq C$, then $c(C,A) = N(C, C \cap A) = N(C,A)$, $c(B,A) = N(B, B \cap A) = N(B,A)$ and $c(C,B) = N(C, C \cap B) = N(C,B)$. Since $N(C,A) \leq N(B,A)$, we have $c(C,A) \leq c(B,A)$. Since $N(C,A) \leq N(C,B)$, we have $c(C,A) \leq c(C,B)$.

(2) **(C1)** The sufficiency has been proved in (1), we only prove the necessity here. If $c(A,B) = 1$, then $N(A, A \cap B) = 1$. By N4 we have $A = A \cap B$. Thus $A \subseteq B$ holds. **(C2)** If $c(A, A^c) = 0$, then $N(A, A \cap A^c) = 0$. As $N$ satisfies N8 and $[\frac{1}{2}] \subseteq A$ we have $A = X$. On the contrary, if $A = X$, then $c(A, A^c) = N(A, A \cap A^c) = N(X, \emptyset) = 0$. **(C3)** We only prove that $A \subseteq B$ implies $c(C,A) \leq c(C,B)$ here. If $A \subseteq B$, then $C \cap A \subseteq C \cap B \subseteq C$, we have $c(C,A) = N(C, C \cap A)$, $c(C,B) = N(C, C \cap B)$. Since $N(C, C \cap A) \leq N(C, C \cap B)$, then $c(C,A) \leq c(C,B)$.

(3) It follows directly from (2).

(4) We only prove that the necessity of C2 also holds when $N$ satisfies N5. If $[\frac{1}{2}] \subseteq A$, then $c(A, A^c) = N(A, A \cap A^c) = N(A, A^c) = 0$. As N satisfies N5 we have $A \cap A^c = \emptyset$, that is $A^c = \emptyset$, thus $A = X$.

(5) It follows directly from (4).

(6) It can be proved in the same manner with (2). □

**Example 14.** Consider the following two similarity measures:

$$N_2(A, B) = \frac{1}{n} \sum_{i=1}^{n} \frac{\min(A(x_i), B(x_i))}{\max(A(x_i), B(x_i))}.$$

$$N_{14}(A, B) = \frac{\sum_{i=1}^{n} \min(A(x_i), B(x_i))}{\sum_{i=1}^{n} \max(A(x_i), B(x_i))}.$$

It is shown that $N_2$ satisfies properties N4, N6, N7, N8, N9 and $N_{14}$ satisfies properties N4, N5, N6. By Proposition 29, we obtain the following subsethood measures:

$$c_1(A, B) = \frac{1}{n} \sum_{i=1}^{n} \frac{\min(A(x_i), B(x_i))}{A(x_i)}.$$

$$c_2(A, B) = \frac{\sum_{i=1}^{n} \min(A(x_i), B(x_i))}{\sum_{i=1}^{n} A(x_i)}.$$

We can conclude that $c_1$ is a *DI*-subsethood measure and it is also a *VY*-subsethood measure and $*$-subsethood measure. We also conclude that $c_2$ is a *VY*-subsethood measure and $*$-subsethood measure.

**Proposition 30.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $N$ be a similarity measure, $c$ a function defined for all $A, B \in F(X)$ by $c(A, B) = N(A \cup B, B)$. Then we obtain the following conclusions:*

(1) $c$ is a subsethood measure,

(2) if $N$ satisfies N4, N6 and N9, then $c$ is a *VY*-subsethood measure,

(3) if $N$ satisfies N4 and N9, then $c$ is a $*$-subsethood measure,

(4) if $N$ satisfies N4, N5 and N6, then $c$ is a *VY*-subsethood measure,

(5) if $N$ satisfies N4 and N5, then $c$ is a $*$-subsethood measure,

(6) if $N$ satisfies N4, N6 and N9, then $c$ is a *DI*-subsethood measure.

*Proof.* It can be proved in the same manner with Proposition 29. □

**Example 15.** Consider the following two similarity measures:

$$N_{12}(A,B)=\frac{\sum_{i=1}^{n}(1-|A(x_i)-B(x_i)|)}{\sum_{i=1}^{n}(1+|A(x_i)-B(x_i)|)}.$$

$$N_{15}(A,B)=\frac{2\sum_{i=1}^{n}\min(A(x_i),B(x_i))}{\sum_{i=1}^{n}(A(x_i)+B(x_i))}.$$

It is shown that $N_{12}$ satisfies properties N4, N6, N7, N8, N9 and $N_{15}$ satisfies properties N4, N5, N6. By Proposition 30, we obtain the following subsethood measures:

$$c_3(A,B)=\frac{\sum_{i=1}^{n}(1-|\max(A(x_i),B(x_i))-B(x_i)|)}{\sum_{i=1}^{n}(1+|\max(A(x_i),B(x_i))-B(x_i)|)}.$$

$$c_4(A,B)=\frac{\sum_{i=1}^{n}2B(x_i)}{\sum_{i=1}^{n}(\max(A(x_i),B(x_i))+B(x_i))}.$$

We can conclude that $c_3$ is a *DI*-subsethood measure and it is also a *VY*-subsethood measure and $*$-subsethood measure. We also conclude that $c_4$ is a *VY*-subsethood measure and $*$-subsethood measure.

**Proposition 31.** *Given a discrete universe $X=\{x_1,x_2,\ldots,x_n\}$. Let $N$ be a similarity measure, $c$ a function defined for all $A,B\in F(X)$ by $c(A,B)=N(A^c,A^c\cup B^c)$. Then we obtain the following conclusions:*

(1) $c$ is a subsethood measure,

(2) if $N$ satisfies N4 and N9, then $c$ is a *VY*-subsethood measure,

(3) if $N$ satisfies N4 and N9, then $c$ is a $*$-subsethood measure,

(4) if $N$ satisfies N4 and N5, then $c$ is a *VY*-subsethood measure,

(5) if $N$ satisfies N4 and N5, then $c$ is a $*$-subsethood measure,

(6) if $N$ satisfies N4, N6 and N9, then $c$ is a *DI*-subsethood measure.

*Proof.* It can be proved in the same manner with Proposition 29. □

**Proposition 32.** *Given a discrete universe $X=\{x_1,x_2,\ldots,x_n\}$. Let $N$ be a similarity measure, $c$ a function defined for all $A,B\in F(X)$ by $c(A,B)=N(A^c\cap B^c,B^c)$. Then we obtain the following conclusions:*

(1) $c$ is a subsethood measure.

(2) if $N$ satisfies N4, N7 and N8, then $c$ is a *VY*-subsethood measure.

(3) if $N$ satisfies N4 and N8, then $c$ is a $*$-subsethood measure.

(4) if $N$ satisfies N4, N5 and N7, then $c$ is a *VY*-subsethood measure.

(5) if $N$ satisfies N4 and N5, then $c$ is a $*$-subsethood measure.

(6) if $N$ satisfies N4, N7 and N8, then $c$ is a *DI*-subsethood measure.

*Proof.* Proof. It can be proved in the same manner with Proposition 29. □

In the following proposition we derive similarity measures from subsethood measures.

**Proposition 33.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $c$ be a subsethood measure, $T$ a t-norm, $N$ a function defined for all $A, B \in F(X)$ by $N(A,B) = T(c(A,B), c(B,A))$, then $N$ is a similarity measure.*

*Proof.*
(N1) $N(X, \emptyset) = T(c(X, \emptyset), c(\emptyset, X)) = 0$ and $N(A,A) = T(c(A,A), c(A,A)) = 1$.
(N2) $N(A,B) = T(c(A,B), c(B,A)) = T(c(B,A), c(A,B)) = N(B,A)$.
(N3) If $A \subseteq B \subseteq C$, then $N(A,C) = T(c(A,C), c(C,A)) = c(C,A)$, $N(A,B) = T(c(A,B), c(B,A)) = c(B,A)$ and $N(B,C) = T(c(B,C), c(C,B)) = c(C,B)$. Since $c(C,A) \leq c(B,A)$ and $c(C,A) \leq c(C,B)$, then $N(A,C) \leq N(A,B)$ and $N(A,C) \leq N(B,C)$. □

**Corollary 1.** *In the conditions of Proposition 33, let $T = T_P$, then the similarity measure derived by subsethood measure can be expressed as:*

$$N(A,B) = c(A,B) \cdot c(B,A).$$

**Corollary 2.** *In the conditions of Proposition 33, let $T = T_M$, then the similarity measure derived by subsethood measure can be expressed as:*

$$N(A,B) = \min(c(A,B), c(B,A)).$$

*Remark* 13. We see that the two formulae given in Corollaries 1 and 2 are the same as the ones given by Zeng and Li [44]. Hence Zeng's solutions can be seen as two special cases of ours.

**Example 16.** Consider the following subsethood measure:

$$c_5(A,B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \frac{\min(1 - A(x_i), 1 - B(x_i))}{1 - B(x_i)}.$$

By Corollary 2, we obtain the following similarity measure:

$$N_{17}(A,B) = \frac{\sum_{i=1}^{n} \min(1 - A(x_i), 1 - B(x_i))}{\sum_{i=1}^{n} \max(1 - A(x_i), 1 - B(x_i))}.$$

## 3.5 Relationships Between Similarity Measures and Fuzzy Entropies

This subsection discusses the relationships between similarity measures and fuzzy entropies and proposes several propositions that similarity measures and fuzzy entropies can be transformed by each other based on their axiomatic definitions.

A measure of fuzzy entropy assesses the amount of vagueness, or fuzziness in a fuzzy set. De Luca and Termini [31] formalize properties of fuzzy entropy through the following axioms.

**Definition 19.** A function $e : F(X) \to [0,1]$ is called an *entropy* on $F(X)$, if $e$ satisfies the following conditions:

(EP1) $e(A) = 0$ if and only if $A$ is nonfuzzy, i.e., $A \in P(X)$,

(EP2) $e(A) = 1$ if and only if $A = [\frac{1}{2}]$,

(EP3) $e(A) \leq e(B)$ in case $A$ refines $B$, i.e., $A(x) \leq B(x)$ when $B(x) \leq \frac{1}{2}$ and $A(x) \geq B(x)$ when $B(x) \geq \frac{1}{2}$.

(EP4) $e(A) = e(A^c)$.

Let $E$ be a fuzzy equivalence, $\mu$ a strictly decreasing function from $[0,1]$ to $[\frac{1}{2}, 1]$ with boundary conditions $\mu(0) = 1$, $\mu(1) = \frac{1}{2}$. For fuzzy sets $A$ and $B$, we define$f(A,B) \in F(X)$, for all $x \in X$,$f(A,B)(x) = \mu(E(A(x), B(x)))$, then we have the following conclusion.

**Proposition 34.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $e$ be a fuzzy entropy, $N$ a function defined for all $A, B \in F(X)$ by $N(A,B) = e(f(A,B))$, then $N$ is a similarity measure.*

*Proof.*
(N1) For all $x \in X$, $f(X, \emptyset)(x) = \mu(E(X(x), \emptyset(x))) = \mu(E(1,0)) = \mu(0) = 1$, then $f(X, \emptyset) = X$. Therefore, $N(X, \emptyset) = e(X) = 0$. Note that $f(A,A)(x) = \mu(E(A(x), A(x))) = \mu(1) = \frac{1}{2}$, then $f(A,A) = [\frac{1}{2}]$. Thus $N(A,A) = e(f(A,A)) = e([\frac{1}{2}]) = 1$.
(N2) It is easy to see that $N(A,B) = e(f(A,B)) = e(f(B,A)) = N(B,A)$.
(N3) $A \subseteq B \subseteq C$ implies $A(x) \leq B(x) \leq C(x)$ for all $x \in X$, by E4 we have $E(A(x), C(x)) \leq \min(E(A(x), B(x)), E(B(x), C(x)))$. By the property of $\mu$, $\mu(E(A(x), C(x))) \geq \max(\mu(E(A(x), B(x)))$, $\mu(E(B(x), C(x)))) \geq \frac{1}{2}$, i.e.,

$$f(A,C)(x) \geq \max(f(A,B)(x), f(B,C)(x)) \geq \tfrac{1}{2}.$$

By EP3, we have $e(f(A,C)) \leq \min(e(f(A,B)), e(f(B,C)))$. Therefore, $N(A,C) \leq \min(N(A,B), N(B,C))$. ☐

*Remark* 14. In Proposition 34, let $\mu(x) = 1 - \frac{1}{2}x, E(x,y) = 1 - |x-y|^n$,then we have $f(A,B)(x) = \frac{1+|A(x)-B(x)|^n}{2}$. Thus the similarity measure $e(f(A,B))$ is in accord with the one given in [44].

**Example 17.** Consider the fuzzy equivalence $E(x,y) = \varphi^{-1}(1 - |\varphi(x) - \varphi(y)|)$, where $\varphi$ is an automorphism of the unit interval. Suppose $\mu(x) = 1 - \frac{1}{2}x$, then we have $f(A,B)(x) = 1 - \frac{1}{2}\varphi^{-1}(1 - |\varphi(A(x)) - \varphi(B(x))|)$. Consider the following fuzzy entropy:

$$e_1(A) = \frac{2}{n} \sum\nolimits_{i=1}^{n} \min(A(x_i), 1 - A(x_i)).$$

By Proposition 34, we obtain the following similarity measure:

$$N_{19}(A,B) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \varphi^{-1}(1 - |\varphi(A(x_i)) - \varphi(B(x_i))|).$$

Note that $N_{19}$ is a similarity measure constructed in [8].

Let $E$ be a fuzzy equivalence, $\lambda$ a strictly increasing function from $[0,1]$ to $[0,\frac{1}{2}]$ with boundary conditions $\lambda(0)=0$, $\lambda(1)=\frac{1}{2}$. For fuzzy sets $A$ and $B$, we define $g(A,B)\in F(X)$, for all $x\in X$, $g(A,B)(x)=\lambda(E(A(x),B(x)))$, then we have the following conclusion.

**Proposition 35.** *Given a discrete universe $X=\{x_1,x_2,\ldots,x_n\}$. Let $e$ be a fuzzy entropy, $N$ a function defined for all $A,B\in F(X)$ by $N(A,B)=e(g(A,B))$, then $N$ is a similarity measure.*

*Proof.* It can be proved in the same manner with Proposition 34. □

Let $E$ be a fuzzy equivalence satisfying E5 and E9, $\mu$ and $\lambda$ the functions defined as above. For fuzzy set $A$, we define $p(A)$, $q(A)\in F(X)$, for all $x\in X$, $p(A)(x)=f(A,A^c)(x)=\mu(E(A(x),A^c(x)))$, $q(A)(x)=g(A,A^c)(x)=\lambda(E(A(x),A^c(x)))$. By the definitions of $\mu$ and $\lambda$ we know that $q(A)(x)\le p(A)(x)$ for all $x\in X$, i.e., $q(A)\subseteq p(A)$. We have the following conclusion.

**Proposition 36.** *Given a discrete universe $X=\{x_1,x_2,\ldots,x_n\}$. Let $N$ be a similarity measure satisfying N5 and N6, $e$ a function defined for all $A\in F(X)$ by $e(A)=N(p(A),q(A))$, then $e$ is a fuzzy entropy.*

*Proof.*
(EP1) (**Necessity**) If $N(p(A),q(A))=e(A)=0$, then by N5, $p(A)\cap q(A)=\emptyset$. Since $q(A)\subseteq p(A)$, we have $q(A)(x)=\lambda(E(A(x),A^c(x)))=0$ for all $x\in X$. Since $\lambda$ is a strictly increasing function satisfying $\lambda(0)=0$, then $E(A(x),A^c(x))=0$. As $E$ satisfies E9 we have $A(x)=1$ or $A(x)=0$. Therefore, $A$ is nonfuzzy.
(**Sufficiency**) If $A$ is nonfuzzy, then we have $A(x)=1$ or $A(x)=0$ for all $x\in X$. Thus $E(A(x),A^c(x))=0$. This means that $q(A)(x)=\lambda(0)=0$, $p(A)(x)=\mu(0)=1$ for all $x\in X$. Therefore, $q(A)=\emptyset$, $p(A)=X$. We have $e(A)=N(p(A),q(A))=N(X,\emptyset)=0$.
(EP2) (**Necessity**) If $N(p(A),q(A))=e(A)=1$, then by N6, $p(A)=q(A)$. Therefore, $p(A)(x)=\mu(E(A(x),A^c(x)))=\lambda(E(A(x),A^c(x)))=q(A)(x)$ for all $x\in X$. By the properties of $\mu$ and $\lambda$, we have $E(A(x),A^c(x))=1$. As $E$ satisfies E9 we have $A(x)=A^c(x)$, that is to say, $A(x)=\frac{1}{2}$.
(**Sufficiency**) If $A=[\frac{1}{2}]$, then $A(x)=\frac{1}{2}$ for all $x\in X$. Thus we conclude that $E(A(x),A^c(x))=E(\frac{1}{2},\frac{1}{2})=1$. This means that $q(A)(x)=\lambda(1)=\frac{1}{2}$, $p(A)(x)=\mu(1)=\frac{1}{2}$ for all $x\in X$. Therefore, we have $p(A)=q(A)$, $e(A)=N(p(A),q(A))=1$.
(EP3) For $x\in X$, if $A(x)\ge B(x)\ge\frac{1}{2}$, then $A^c(x)\le B^c(x)\le\frac{1}{2}$, thus $A^c(x)\le B^c(x)\le\frac{1}{2}\le B(x)\le A(x)$. By E4 we have $E(A(x),A^c(x))\le E(B(x),B^c(x))$. By the properties of $\mu$ and $\lambda$, $\lambda(E(A(x),A^c(x)))\le\lambda(E(B(x),B^c(x)))\le\mu(E(B(x),B^c(x)))\le\mu(E(A(x),A^c(x)))$. This means that $q(A)\subseteq q(B)\subseteq p(B)\subseteq p(A)$. Thus $N(p(A),q(A))\le N(p(B),q(A))\le N(p(B),q(B))$. That is to say, $e(A)\le e(B)$. The case of $A(x)\le B(x)\le\frac{1}{2}$ can be proved similarly.
(EP4) Since $p(A)=p(A^c)$, $q(A)=q(A^c)$, then $e(A)=N(p(A),q(A))=N(p(A^c),q(A^c))=e(A^c)$. □

*Remark* 15.
(1) In conditions of Proposition 36, let $\mu(x)=1-\frac{1}{2}x$, $\lambda(x)=\frac{1}{2}x$, $E(x,y)=$

$1-|x-y|^n$, then we have $p(A)(x) = \frac{1+|A(x)-A^c(x)|^n}{2}$, $q(A)(x) = \frac{1-|A(x)-A^c(x)|^n}{2}$. Therefore, the fuzzy entropy $N(p(A), q(A))$ is in accord with the one given by Zeng and Li [44].
(2) In conditions of (1), if $n = 1$, then we have $p(A)(x) = \frac{1+|A(x)-A^c(x)|}{2} = \max(A(x), A^c(x))$ and $q(A)(x) = \frac{1-|A(x)-A^c(x)|}{2} = \min(A(x), A^c(x))$. This means that $p(A) = A \cup A^c$, $q(A) = A \cap A^c$. Thus the fuzzy entropy constructed by similarity measure $N$ can be expressed as $N(A \cup A^c, A \cap A^c)$. This is the solution given by Fan [13].

**Example 18.** Suppose $\mu(x) = 1-\frac{1}{2}x$, $\lambda(x) = \frac{1}{2}x$, $E(x, y) = \varphi^{-1}(1-|\varphi(x)-\varphi(y)|)$, where $\varphi$ is an automorphism of the unit interval. Since $E$ is a fuzzy equivalence satisfying E5 and E9, we have

$$q(A)(x) = \lambda(E(A(x), A^c(x))) = \frac{1}{2}\varphi^{-1}(1-|\varphi(A(x))-\varphi(1-A(x))|),$$

$$p(A)(x) = \mu(E(A(x), A^c(x))) = 1-\frac{1}{2}\varphi^{-1}(1-|\varphi(A(x))-\varphi(1-A(x))|).$$

Consider the following similarity measure:

$$N_{14}(A, B) = \frac{\sum_{i=1}^{n}\min(A(x_i), B(x_i))}{\sum_{i=1}^{n}\max(A(x_i), B(x_i))}.$$

It is shown that $N_{14}$ satisfies properties N4, N5, N6. By Proposition 36, we obtain the following fuzzy entropy:

$$e_2(A) = \frac{\sum_{i=1}^{n}\varphi^{-1}(1-|\varphi(A(x_i))-\varphi(1-A(x_i))|)}{\sum_{i=1}^{n}(2-\varphi^{-1}(1-|\varphi(A(x_i))-\varphi(1-A(x_i))|))}.$$

**Proposition 37.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $N$ be a similarity measure satisfying N5 and N6, $e$ a function defined for all $A \in F(X)$ by $e(A) = N(A, A^c)$, then $e$ is a fuzzy entropy.*

*Proof.* It follows from Proposition 36 that this proposition holds. □

## 3.6 Relationships Between Subsethood Measures and Fuzzy Entropies

This subsection discusses the relationships between subsethood measures and fuzzy entropies and proposes several propositions that subsethood measures and fuzzy entropies can be transformed by each other based on their axiomatic definitions.

At first blush, subsethood measure and fuzzy entropy do not seem related. To relate subsethood measure with fuzzy entropy, Kosko [23] proposed the following expression: given a subsethood measure $c$, the fuzzy entropy $e$ generated by $c$ is defined as $e(A) = c(A \cup A^c, A \cap A^c)$ for all $A \in F(X)$. For showing the conditions of $c$ for which $e$ can be a fuzzy entropy, several axiomatizations were given in the literature. In the previous part of this chapter, we have referred to three concepts of subsethood measure, that is, *VY*-subsethood measure, $*$-subsethood measure, and *DI*-subsethood measure. It is shown that all *DI*-subsethood measure is *VY*-subsethood measure and therefore, it is also $*$-subsethood measure. We know that the three conditions of $*$-subsethood measure are enough to demand the conditions for the expression: $e(A) = c(A \cup A^c, A \cap A^c)$ to fulfill the conditions demanded from fuzzy entropy. Therefore, we use $*$-subsethood measure to construct fuzzy entropy here.

**Proposition 38.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $c$ be a $*$-subsethood measure, $p(A)$ defined as above. Suppose $(p(A))^c$ is the complement of the fuzzy set $p(A)$. If $e$ is a function defined for all $A \in F(X)$ by $e(A) = c(p(A), (p(A))^c)$, then $e$ is a fuzzy entropy.*

*Proof.*
(EP1) (**Necessity**) Since $p(A)(x) \geq \frac{1}{2}$, then we have $[\frac{1}{2}] \subseteq p(A)$. If $c(p(A), (p(A))^c) = 0$, then by C2 of $*$-subsethood measure, we have $p(A) = X$. This means that $p(A)(x) = \mu(E(A(x), A^c(x))) = 1$ for all $x \in X$. Thus according to the properties of $\mu$ we have $E(A(x), A^c(x)) = 0$. As $E$ satisfies E9 we have $A(x) = 1$ or $A(x) = 0$. Therefore, $A$ is nonfuzzy.
(**Sufficiency**) $A$ is nonfuzzy implies $A(x) = 1$ or $A(x) = 0$ for all $x \in X$. This means that $E(A(x), A^c(x)) = 0$. Thus $p(A)(x) = \mu(0) = 1$, $p(A^c)(x) = 0$, that is to say, $p(A) = X$. Therefore, we have $e(A) = c(p(A), (p(A))^c) = c(X, \emptyset) = 0$.
(EP2) (**Necessity**) If $c(p(A), (p(A))^c) = e(A) = 1$, then by C1 of $*$-subsethood measure, $p(A) \subseteq (p(A))^c$. As $(p(A))^c \subseteq p(A)$ we have $(p(A))^c = p(A)$. Thus $\mu(E(A(x), A^c(x))) = \frac{1}{2}$for all $x \in X$. According to properties of $\mu$, we have $E(A(x), A^c(x)) = 1$. As $E$ satisfies E5 we have $A(x) = A^c(x)$, that is to say, $A(x) = \frac{1}{2}$.
(**Sufficiency**) If $A = [\frac{1}{2}]$, then $A(x) = \frac{1}{2}$ for all $x \in X$. Thus we conclude that $E(A(x), A^c(x)) = 1$. This means that $p(A)(x) = (p(A))^c(x) = \frac{1}{2}$, i.e., $(p(A))^c = p(A)$. Therefore, $e(A) = c(p(A), (p(A))^c) = 1$.
(EP3) For $x \in X$, if $A(x) \leq B(x) \leq \frac{1}{2}$, then $A^c(x) \geq B^c(x) \geq \frac{1}{2}$, thus $A(x) \leq B(x) \leq \frac{1}{2} \leq B^c(x) \leq A^c(x)$. By E4 we have $E(A(x), A^c(x)) \leq E(B(x), B^c(x))$. By the properties of $\mu$, we have $1 - \mu(E(A(x), A^c(x))) \leq 1 - \mu(E(B(x), B^c(x))) \leq \mu(E(B(x), B^c(x))) \leq \mu(E(A(x), A^c(x)))$. This means that $(p(A))^c \subseteq (p(B))^c \subseteq p(B) \subseteq p(A)$. By C3 of $*$-subsethood measure, we have $c(p(A), (p(A))^c) \leq c(p(B), (p(A))^c) \leq c(p(B), (p(B))^c)$. That is to say, $e(A) \leq e(B)$. The case of $A(x) \geq B(x) \geq \frac{1}{2}$ can be proved similarly.
(EP4) Since $p(A) = p(A^c)$ and $q(A) = q(A^c)$, then $e(A) = c(p(A), (p(A))^c) = c(p(A^c), (p(A^c))^c) = e(A^c)$. □

*Remark* 16. In conditions of Proposition 38, let $\mu(x) = 1 - \frac{1}{2}x$, $E(x, y) = 1 - |x - y|$, then we have

$$p(A)(x) = \tfrac{1+|A(x)-A^c(x)|}{2} = \max(A(x), A^c(x)),$$

$$(p(A))^c(x) = \tfrac{1-|A(x)-A^c(x)|}{2} = \min(A(x), A^c(x)).$$

It means that $p(A) = A \cup A^c$, $p(A))^c = A \cap A^c$. Thus the fuzzy entropy derived from subsethood measure $c$ can be expressed as $c(A \cup A^c, A \cap A^c)$. In this sense, Kosko's solution [23] about the relation between fuzzy entropy and subsethood measure can be brought into line with our solution.

**Example 19.** Suppose $\mu(x) = 1 - \frac{1}{2}x$, $E(x, y) = \varphi^{-1}(\frac{\min(\varphi(x),\varphi(y))}{\max(\varphi(x),\varphi(y))})$, where $\varphi$ is an automorphism of the unit interval. Since $E$ is a fuzzy equivalence satisfying E5 and E9, we have

$$p(A)(x) = \mu(E(A(x), A^c(x))) = 1 - \frac{1}{2}\varphi^{-1}(\frac{\min(\varphi(A(x)), \varphi(1 - A(x)))}{\max(\varphi(A(x)), \varphi(1 - A(x)))}),$$

$$p(A))^c(x) = 1 - \mu(E(A(x), A^c(x))) = \frac{1}{2}\varphi^{-1}(\frac{\min(\varphi(A(x)), \varphi(1 - A(x)))}{\max(\varphi(A(x)), \varphi(1 - A(x)))}).$$

Consider the following subsethood measure:

$$c_6(A, B) = \frac{\sum_{i=1}^{n} B(x_i)}{\sum_{i=1}^{n} \max(A(x_i), B(x_i))}.$$

It is shown that $c_6$ is a $*$-subsethood measure. By Proposition 38, we obtain the following fuzzy entropy:

$$e_3(A) = \frac{\sum_{i=1}^{n} \varphi^{-1}(\frac{\min(\varphi(A(x_i)),\varphi(1-A(x_i)))}{\max(\varphi(A(x_i)),\varphi(1-A(x_i)))})}{\sum_{i=1}^{n} (2 - \varphi^{-1}(\frac{\min(\varphi(A(x_i)),\varphi(1-A(x_i)))}{\max(\varphi(A(x_i)),\varphi(1-A(x_i)))}))}.$$

Let $E$ be a fuzzy equivalence satisfying E5 and E9, $\mu$ and $\lambda$ the functions defined as above. For fuzzy sets $A$ and $B$, we define $k(A, B), l(A, B) \in F(X)$ for all $x \in X$, $k(A, B)(x) = \mu(E(A(x), \min(A(x), B(x))))$ and $l(A, B)(x) = \lambda(E(A(x), \min(A(x), B(x))))$. Then we have the following conclusion.

**Proposition 39.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $e$ be a fuzzy entropy, $c$ a function defined for all $A, B \in F(X)$ by $c(A, B) = e(k(A, B))$, then $c$ is a DI-subsethood measure.*

*Proof.*
(C1) (**Necessity**) If $c(A, B) = e(k(A, B)) = 1$, by EP2, we have $k(A, B) = [\frac{1}{2}]$. This means that $\mu(E(A(x), \min(A(x), B(x)))) = \frac{1}{2}$ for all $x \in X$. Since $\mu$ is a strictly decreasing function and $\mu(1) = \frac{1}{2}$, then we have $E(A(x), \min(A(x), B(x))) = 1$ for all $x \in X$. As $E$ satisfies E5 we have $A(x) = \min(A(x), B(x))$, thus $A(x) \leq B(x)$ for all $x \leq X$.
(**Sufficiency**) If $A \subseteq B$, then $k(A, B)(x) = \mu(E(A(x), A(x))) = \mu(1) = \frac{1}{2}$, that is to say, $k(A, B) = [\frac{1}{2}]$. Therefore, $c(A, B) = e([\frac{1}{2}]) = 1$.

(C2) (**Necessity**) If $e(k(A, A^c)) = c(A, A^c) = 0$, then by EP1, $k(A, A^c)$ is nonfuzzy. Therefore, $k(A, A^c)(x) = 1$ or $0$. By the definition of $k(A, B)$, we know $k(A, B)(x) \geq \frac{1}{2}$. Thus $k(A, A^c)(x) = 1$ for all $x \in X$. This means that $\mu(E(A(x), \min(A(x), 1 - A(x)))) = 1$. By the properties of $\mu$, we have $E(A(x), \min(A(x), 1 - A(x))) = 0$. If $A(x) \leq \frac{1}{2}$, then $E(A(x), \min(A(x), 1 - A(x))) = E(A(x), A(x)) = 1 \neq 0$. Thus we conclude that $A(x) \geq \frac{1}{2}$for all $x \in X$. Thus $E(A(x), 1 - A(x)) = 0$. As $E$ satisfies

E9 we have $A(x) = 1$ for all $x \in X$.
(**Sufficiency**) If $A = X$, then $k(A, A^c)(x) = \mu(E(1, 0)) = \mu(0) = 1$. Thus $k(A, A^c) = X$. Therefore, $c(A, A^c) = e(k(A, A^c)) = e(X) = 0$.

(C3) Since $A \subseteq B$ implies $A(x) \leq B(x)$ for all $x \in X$, then for any fuzzy set $C$, three cases will be considered depending on the position of $C(x)$.
(1) If $A(x) \leq B(x) \leq C(x)$, then $k(B, C)(x) = \mu(E(B(x), B(x))) = \mu(1) = \frac{1}{2}$, $k(A, C)(x) = \mu(E(A(x), A(x))) = \mu(1) = \frac{1}{2}$ for $x \in X$. Thus $k(B, C)(x) = k(A, C)(x)$.
(2) If $C(x) \leq A(x) \leq B(x)$, then $k(B, C)(x) = \mu(E(B(x), C(x)))$, $k(A, C)(x) = \mu(E(A(x), C(x)))$ for $x \in X$. According to E4, we conclude that $E(B(x), C(x)) \leq E(A(x), C(x))$. As $\mu$ is strictly decreasing we have $\mu(E(B(x), C(x))) \geq \mu(E(A(x), C(x)))$. Thus $k(B, C)(x) \geq k(A, C)(x)$.
(3) If $A(x) \leq C(x) \leq B(x)$, then $k(B, C)(x) = \mu(E(B(x), C(x)))$, $k(A, C)(x) = \mu(E(A(x), A(x)))$ for $x \in X$. Since $E(B(x), C(x)) \leq E(A(x), A(x)) = 1$. As $\mu$ is strictly decreasing we have $\mu(E(B(x), C(x))) \geq \mu(E(A(x), A(x)))$. Thus $k(B, C)(x) \geq k(A, C)(x)$.

Hence we can conclude that $k(B, C)(x) \geq k(A, C)(x) \geq \frac{1}{2}$ for all $x \in X$. Then by EP3 we have $e(k(B, C)) \leq e(k(A, C))$, i.e., $c(B, C) \leq c(A, C)$. The case of $c(C, A) \leq c(C, B)$ whenever $A \subseteq B$ can be proved similarly. □

**Example 20.** In conditions of Proposition 39, let

$$\mu(x) = 1 - \tfrac{1}{2}x,\ E(x, y) = \varphi^{-1}\left(\frac{\min(\varphi(x), \varphi(y))}{\max(\varphi(x), \varphi(y))}\right),$$

where $\varphi$ is an automorphism of the unit interval, then we can conclude that

$$k(A, B)(x) = 1 - \frac{1}{2}\varphi^{-1}\left(\frac{\varphi(\min(A(x), B(x)))}{\varphi(A(x))}\right).$$

Consider the following fuzzy entropy:

$$e_1(A) = \frac{2}{n}\sum\nolimits_{i=1}^{n} \min(A(x_i), 1 - A(x_i)).$$

By Proposition 39, we obtain the following *DI*-subsethood measure:

$$c_7(A, B) = \frac{1}{n}\sum\nolimits_{i=1}^{n} \varphi^{-1}\left(\frac{\varphi(\min(A(x_i), B(x_i)))}{\varphi(A(x_i))}\right).$$

**Proposition 40.** *Given a discrete universe $X = \{x_1, x_2, \ldots, x_n\}$. Let $e$ be a fuzzy entropy, $c$ a function defined for all $A, B \in F(X)$ by $c(A, B) = e(l(A, B))$, then $c$ is a DI-subsethood measure.*

*Proof.* It can be proved in the same manner with Proposition 39. □

**Corollary 3.** *Note that the subsethood measures constructed by fuzzy entropy in the above-mentioned propositions are also VY-subsethood measures. Therefore, they are also $*$-subsethood measures.*

## Acknowledgement

This research is supported by the National Natural Science Foundation of China (Grant Nos. 61603307, 61305074, 61473239 and 71301129), the Fundamental Research Funds for the Central Universities of China (Grant Nos. JBK160132, A0920502051603-52 and JBK 150503) and the Grant from MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Grant No. 12YJCZH155).


## References


[1] C. Bertoluzza, P. Miranda, and P. Gil. A generalization of local divergence measures. *International Journal of Approximate Reasoning*, 40(3):127–146, 2005.

[2] D. Bhandari and N.R. Pal. Some new information measures for fuzzy sets. *Information Sciences*, 67(3):209–228, 1993.

[3] I. Bloch. On fuzzy distances and their use in image processing under imprecision. *Pattern Recognition*, 32(11):1873–1895, 1999.

[4] B. Bouchon-Meunier, M. Rifqi, and S. Bothorel. Towards general measures of comparison of objects. *Fuzzy Sets and Systems*, 84(2):143–153, 1996.

[5] J.J. Buckley and Y. Hayashi. Fuzzy input-output controllers are universal approximators. *Fuzzy Sets and Systems*, 58(3):273–278, 1993.

[6] H. Bustince, E. Barrenechea, and M. Pagola. Restricted equivalence functions. *Fuzzy Sets and Systems*, 157(17):2333–2346, 2006.

[7] H. Bustince, E. Barrenechea, and M. Pagola. Image thresholding using restricted equivalence functions and maximizing the measures of similarity. *Fuzzy Sets and Systems*, 158(5):496–516, 2007.

[8] H. Bustince, E. Barrenechea, and M. Pagola. Relationship between restricted dissimilarity functions, restricted equivalence functions and normal en-functions: Image thresholding invariant. *Pattern Recognition Letters*, 29(4):525–536, 2008.

[9] H. Bustince, V. Mohedano, and E. Barrenechea. Definition and construction of fuzzy di-subsethood measures. *Information Sciences*, 176(21):3190–3231, 2006.

[10] S.M. Chen, M.S. Yeh, and P.Y. Hsiao. A comparison of similarity measures of fuzzy values. *Fuzzy Sets and Systems*, 72(1):79–89, 1995.

[11] I. Couso, V. Janiš, and S. Montes. Fuzzy divergence measures. *Acta Universitatis Matthiae Belii series Mathematics*, 8:21–26, 2000.

[12] J. Diatta. Description-meet compatible multiway dissimilarities. *Discrete Applied Mathematics*, 154(3):493–507, 2006.

[13] J. Fan. *Subsethood Measures*. Northwestern Press, Hsian, 1999. (in Chinese).

[14] J. Fan. Some new similarity measures. *Journal of Xian Institute of Posts and Telecommunications*, 7(3):69–71, 2002. (in Chinese).

[15] J. Fan and W. Xie. Distance measure and induced fuzzy entropy. *Fuzzy Sets and Systems*, 104(2):305–314, 1999.

[16] J. Fan and W. Xie. Some notes on similarity measure and proximity measure. *Fuzzy Sets and Systems*, 101(3):403–412, 1999.

[17] J. Fan, W. Xie, and J. Pei. Subsethood measure: new definitions. *Fuzzy Sets and Systems*, 106(2):201–209, 1999.

[18] J.C. Fodor and M.R. Roubens. *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer Academic Publishers, Dordrecht, 1994.

[19] H. Guoshun and L. Yunsheng. New subsethood measures and similarity measures of fuzzy sets. In *International Conference on Communications, Circuits and Systems (ICCCAS)*, pages 999–1002, 2005.

[20] D.H. Hong and S.Y. Hwang. A note on the value similarity of fuzzy systems variables. *Fuzzy Sets and Systems*, 66(3):383–386, 1994.

[21] E.P. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Kluwer Academic Publishers, Dordrecht, 2000.

[22] A. Kolesárová. Limit properties of quasi-arithmetic means. *Fuzzy Sets and Systems*, 124(1):65–71, 2001.

[23] B. Kosko. Fuzzy entropy and conditioning. *Information Sciences*, 40(2):165–174, 1986.

[24] B. Kosko. Fuzziness vs. probability. *International Journal of General Systems*, 17(2-3):211–240, 1990.

[25] H. Lee-Kwang, Y.S. Song, and K.M. Lee. Similarity measure between fuzzy sets and between elements. *Fuzzy Sets and Systems*, 62(3):291–293, 1994.

[26] X. Li and B. Liu. On distance between fuzzy variables. *Journal of Intelligent & Fuzzy Systems*, 19(3):197–204, 2008.

[27] Y. Li, K. Qin, and X. He. Relations among similarity measure, subsethood measure and fuzzy entropy. *International Journal of Computational Intelligence Systems*, 6(3):411–422, 2013.

[28] Y. Li, K. Qin, and X. He. Dissimilarity functions and divergence measures between fuzzy sets. *Information Sciences*, 288:15–26, 2014.

[29] Y. Li, K. Qin, and X. He. Some new approaches to constructing similarity measures. *Fuzzy Sets and Systems*, 234:46–60, 2014.

[30] Y. Li, K. Qin, X. He, and D. Meng. Similarity measures of interval-valued fuzzy sets. *Journal of Intelligent & Fuzzy Systems*, 28(5):2113–2125, 2015.

[31] A. De Luca and S. Termini. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Information and Control*, 20(4):301–312, 1972.

[32] S. Montes, I. Couso, P. Gil, and C. Bertoluzza. Divergence measure between fuzzy sets. *International Journal of Approximate Reasoning*, 30(2):91–105, 2002.

[33] S. Montes, P. Miranda, J. Jiménez, and P. Gil. A measure of the difference between two fuzzy paptitions. *Tatra Mountains Mathematical Publications*, 21:133–152, 2001.

[34] C.P. Pappis. Value approximation of fuzzy systems variables. *Fuzzy Sets and Systems*, 39(1):111–115, 1991.

[35] C.P. Pappis and N.I. Karacapilidis. A comparative assessment of measures of similarity of fuzzy values. *Fuzzy Sets and Systems*, 56(2):171–174, 1993.

[36] J. Recasens. *Indistinguishability Operators: Modelling Fuzzy Equalities and Fuzzy Equivalence Relations*, volume 260 of *Studies in Fuzziness and Soft Computing*. Springer-Verlag, Heidelberg, 2011.

[37] I.B. Turksen and Z. Zhong. An approximate analogical reasoning approach based on similarity measures. *IEEE Transactions on Systems, Man, and Cybernetics*, 18(6):1049–1056, 1988.

[38] P.Z. Wang. *Fuzzy Sets and Its Applications*. Shanghai Science and Technology

Press, Shanghai, 1983. (in Chinese).

[39] W.J. Wang. New similarity measures on fuzzy sets and on elements. *Fuzzy Sets and Systems*, 85(3):305–309, 1997.

[40] X. Wang, B. De Baets, and E. Kerre. A comparative study of similarity measures. *Fuzzy Sets and Systems*, 73(2):259–268, 1995.

[41] X. Wang, D. Ruan, and E.E. Kerre. *Mathematics of Fuzziness-Basic Issues*, volume 245 of *Studies in Fuzziness and Soft Computing*. Springer-Verlag, Heidelberg, 2009.

[42] R. Willmott. On the transitivity of containment and equivalence in fuzzy power set theory. *Journal of Mathematical Analysis and Applications*, 120(1):384–396, 1986.

[43] L. Xuecheng. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems*, 52(3):305–318, 1992.

[44] W Zeng and H. Li. Inclusion measures, similarity measures, and the fuzziness of fuzzy sets and their relations. *International Journal of Intelligent Systems*, 21(6):639–653, 2006.