II YEAR – III SEMESTER
COURSE CODE: 7MMA3E4

ELECTIVE COURSE- IV- (A) – FUZZY MATHEMATICS

Unit I

Crisp sets and fuzzy sets.

Unit II

Operation on fuzzy sets.

Unit III

Fuzzy relations.

Unit IV

Fuzzy measures.

Unit V

Uncertainty and Information.

Text Books

1. J.Klir and Tina A Folger, Fizzy Sets, Uncertainty and Information, Prentice Hall of India Private Ltd., New Delhi, 2006

Chapters : I, II, III, IV and V upto section 5.5.

Books for Supplementary Reading and Reference:

1. V.Novak, Fuzzy Sets and Their Applications, Adom Hilger, Bristol, 1969.
2. A.Kaufman, Introduction to the Theory of Fuzzy Subsets, Academic Press, 1975.
3. H.J.Zimmermann, Fuzzy Set Theory and its Applications, Allied Publishers, Chennai, 1996.

Notations:

Z = set of all integers
$= \{ \ldots, -2, -1, 0, 1, 2, \ldots \}$

N = set of all positive integers
(or)
set of all natural numbers
$= \{1, 2, \ldots\}$

W = set of all non-negative integers
(or)
set of all whole numbers

$= \{0, 1, 2, \ldots\}$

$N_n = \{1, 2, \ldots n\}$

$W_n = \{0, 1, 2, \ldots n\}$

$Z_n = \{0, 1, 2, \ldots n-1\}$

R = the set of all real numbers

R^+ = the set of all non-negative real numbers.

Fuzzy set (or) Fuzzy subset:

Defn 1.1: A fuzzy subset A on a universal set X is a function. $A : X \to [0,1]$ where $[0,1] \subset R$

Remark: Another forms of the fuzzy subset are

(i) $A = \{ (x, \mu_A(x)) / x \in X \}$, where $\mu_A : X \to [0,1]$ is called a membership function of A.

(ii) $A = \sum_{i=1}^{n} A(x_i)/x_i$ (Discrete values)

(iii) $A = \int_X A(x)/x$, (Interval values)

Example: 1) Let $X = \{1, 2, 3, 4, 5\}$. Then the fuzzy subset A of X is $A = \{ (1, .3), (2, .6), (3, .7), (4, .5), (5, .1) \}$.

1.2) A = "real number close to 0". Then we can define the fuzzy subset A as
$A = \{(x, A(x)) / A(x) = (1+10x^2)^{-1}\}$.

1.3) A = "real number very close to 0". Then we can define the fuzzy subset A as
$A = \{x, A(x)) / A(x) = (1+10x^2)^{-2}\}$

1.4) A = "real numbers close to any given number a". Thus we can define the fuzzy subset A as $A = \{(x, A(x)) / A(x) = (1+10(x-a)^2)^{-1}$.

1.5) A = " Integers close to 10". Then we can define the fuzzy subset A as $A = \{ \frac{0.1}{7} + \frac{0.5}{8} + \frac{0.8}{9} + \frac{1}{10} + \frac{0.8}{11} + \frac{0.5}{12} + \frac{0.1}{13}$

5m ✓ Types of fuzzy Subsets :

L fuzzy Subset 1.7 :

A fuzzy subset A on a universal set X is a function $A : X \to L$ where L is a Lattice. Then the fuzzy subset A is called a L-Fuzzy subset of X.

Example's 1.8 :

1) $X = \{1, 2, 3, 4\}$. A function $A : X \to L$ is defined as $A(1) = 0.8$; $A(2) = 0.2$, $A(3) = 0.1$, $A(4) = 0.6$

The L - Fuzzy subset A is

$A = \{(1, 0.8), (2, 0.2), (3, 0.1), (4, 0.6)\}$.

Interval valued fuzzy Subset 1.9 :

A fuzzy subset A on a universal of X is a function $A : X \to D([0,1])$ where $D([0,1])$ is a family of all closed intervals of real numbers in $[0,1]$. Then the fuzzy subset is called a interval valued fuzzy subset of X.

Example 1.10

Let $X = \{1, 2, 3\}$ The interval valued fuzzy subset A of X is defined as

$A = \{(1, [0.3, 0.4]), (2, [0.1, 0.6]), (3, [0.6, 0.$

Fuzzy subsets of Type 2 1.11.

A fuzzy subset A on a universal set X is a function $A: X \to F([0,1])$, where $F([0,1])$ is the set of all ordinary fuzzy subsets that can be defined within the interval $[0,1]$. (it is called a fuzzy power subset of $[0,1]$). The fuzzy subset A is called a fuzzy subset of type 2.

Example 1.12: Let $X = \{1, 2, 3\}$ The fuzzy subset of type 2 A is defined as

$A = \{(1, (.4, .8)), (2, (0.6, .9)), \cancel{(3, (.7,}$

$A = \{(1, A_1), (2, A_2), (3, A_3)\}$ where $A_1, A_2, \ldots F([0,1])$

Fuzzy measure 1.13:

A fuzzy subset A on a power set $P(X)$ is a function $A: P(X) \to [0,1]$. The fuzzy subset A of $P(X)$ is called a fuzzy measure.

Example: 1.14: Let $X = \{1, 2, 3\}$. The fuzzy measure A of $P(X)$ is defined as

$A = \{(X, .3), (\{1, 2\}, .4), (\{1\}, .6), (\{3\}, .5)$

level 2 fuzzy subset 1.15:

A fuzzy subset A on the fuzzy power subset of X is a function $A: F(X) \to [0,1]$ where $F(X)$ is the fuzzy power subset of X.

Complement of fuzzy subset 1.19:

The standard complement $\bar{A}$ of a fuzzy subset A of X is defined as $\bar{A}(x) = 1 - A(x)$, $\forall x \in X$.

Example 1.20:

Let $X = \{5, 10, 20, 30, 40, 50, 60, 70, 80\}$ be a collection of ages.

A = "old age", so the fuzzy subset A of X is $A = 0/5 + 0/10 + .1/20 + .2/30 + \frac{.4}{40} + \frac{.6}{50} + .8/60 + 1/70 + 1/80$

The fuzzy set Not old $\bar{A} = 1/5 + 1/10 + \frac{.9}{20} + \frac{.8}{30} + \frac{.6}{40} + \frac{.4}{50} + \frac{.2}{60} + \frac{0}{70} + \frac{0}{80}$

Standard Union 1.21:

Let A and B be any two fuzzy subsets of X. Then their union $A \cup B$ is defined as

$(A \cup B)(x) = \max\{A(x), B(x)\}$, $\forall x \in X$.

Example 1.22: Let $X = \{5, 10, 20, 30, 40, 50, 60, 70, 80\}$ be a collection of ages. A = "young ages". So the fuzzy subset A of X is $A = 1/5 + 1/10 + \frac{0.8}{20} + \frac{.5}{30} + \frac{.2}{40} + \frac{.1}{50} + \frac{0}{60} + \frac{0}{70} + \frac{0}{80}$, and B = "old ages", so the fuzzy subset B of X is $B = \frac{0}{5} + \frac{0}{10} + \frac{.1}{20} + \frac{.2}{30} + \frac{.4}{40} + \frac{.6}{50} + \frac{.8}{60} + \frac{1}{70} + \frac{1}{80}$. Then union of A and B is

$A \cup B = \frac{1}{5} + \frac{1}{10} + \frac{.8}{20} + \frac{.5}{30} + \frac{.4}{40} + \frac{.6}{50} + \frac{.8}{60} + 1/70 + 1/80$

Standard Intersection 1.23:

Let A and B be any two fuzzy subsets of X. Then their intersection $A \cap B$ is defined as $(A \cap B)(x) = \min\{A(x), B(x)\}$, $\forall x \in X$.

$A(x) = B(x), \forall x \in X.$

Not equal 1.34 :

Let A and B be any two fuzzy subsets of X. Thus A and B are called not equal if $A(x) \neq B(x)$, atleast one $x \in X$.

Ex 1.32: Let $A = \{(a, .1), (b, .5), (c, .3)\}$ and $B = \{(a, .2), (b, .4), (c, .8)\}$ be two fuzzy subsets of $X = \{a, b, c\}$. Clearly A and B are not equal, since $A(a) \neq B(a)$.

Defn 1.33: Let A be a fuzzy subset of X. Then the height of A is $h(A) = \sup_{x \in X} A(x)$.

Example 1.33: Let $A = \{(a, .4), (b, .6), (c, .9)\}$ be a fuzzy subset of $X = \{a, b, c\}$. Then the height of A is

$$h(A) = 0.9$$

Defn 1.35: Let A be a fuzzy subset of X. Thus A is called normal if $h(A) = 1$.

Example 1.35: Let $A = \{(a, 1), (b, .4), (c, .8)\}$ be a fuzzy subset of $X = \{a, b, c\}$. Clearly A is normal, since $h(A) = 1$

Defn 1.37: Let A be a fuzzy subsets of X. Then A is called a Subnormal if $h(A) < 1$.

Example 1.38:

Let $A = \{(a, .4), (b, .7), (c, .9)\}$ be a fuzzy subset of $X = \{a, b, c\}$. Clearly A is a subnormal, since $h(A) < 1$.

Defn 1.39: Let A be a fuzzy subset of a finite universal set X. Then its scalar cardinality $|A|$ is defined as $|A| = \sum_{x \in X} A(x)$.

Example 1.40:

Let $X = \{0, 2, 4, 6, \dots 80\}$ be any set. The fuzzy subset A of X is defined as $A(x) = 0$ if $x \notin \{22, 24, \dots 58\}$, $A(x) = 0.13$ if $x \in \{22, 58\}$, $A(x) = 0.27$ if $x \in \{24, 56\}$, $A(x) = 0.4$ if $x \in \{26, 54\}$, $A(x) = 0.53$ if $x \in \{28, 52\}$, $A(x) = 0.67$ if $x \in \{30, 50\}$, $A(x) = 0.8$ if $x \in \{32, 48\}$, $A(x) = 0.93$ if $x \in \{34, 46\}$, $A(x) = 1$ if $x \in \{36, 38, \dots 44\}$.

The scalar cardinality $|A|$ of A is

$$|A| = 2(0.13 + 0.27 + 0.4 + 0.53 + 0.67 + 0.8 + 0.93) + 1 + 1 + 1 + 1 + 1$$

$$= 12.46$$

Note: $|A|$ is also called sigma count of A.

Def: Let A and B be any two fuzzy subsets of a finite universal set X. Thus the degree of subset hood $S(A, B)$ of A in B is defined as $S(A, B) = \frac{1}{|A|}\left(|A| - \sum_{x \in X} \max\{0, A(x) - B(x)\}\right)$

or $S(A, B) = \frac{|A \cap B|}{|A|}$

Note: Clearly [illegible]

1.11 Example: Let $A = \{(a, .3), (b, [illegible]), (c, [illegible]), (d, [illegible])\}$ and $B = \{(a, .2), (b, [illegible]), (c, [illegible]), (d, [illegible])\}$ be two fuzzy subsets of $X = \{a, b, c, d\}$.

Then $A \cap B = \{(a, .2), (b, [illegible]), (c, [illegible]), (d, [illegible])\}$

$\therefore |A \cap B| = .2 + .4 + .3 + .2 = 1.1$

$|A| = .3 + .4 + .6 + .2 = 1.5$

$\therefore S(A, B) = \dfrac{|A \cap B|}{|A|} = \dfrac{1.1}{1.5} = 0.7333..$

1.12 Defn: Let A and B be any two fuzzy subsets of a finite universal set X. Then Hamming distance of A from B is defined as

$$d(A, B) = \sum_{x \in X} |A(x) - B(x)|$$

1.13 Example: Let $A = \{(a, .5), (b, .6), (c, .4), (d, .5)\}$ be two and $B = \{(a, .2), (b, .4), (c, .7), (d, [illegible])\}$ fuzzy subsets of $X = \{a, b, c, d\}$.

Then $d(A, B) = |.5 - .2| + |.6 - .4| + |.4 - .7| + [illegible]$

$= .3 + .2 + .3 + [illegible] = [illegible]$

1.14 Defn: Let A and B be any two fuzzy subsets of a universal set X. Then the difference of A and B is defined as

$$A - B = A \cap \bar{B} = \min\{A(x), 1 - B(x)\}.$$

1.15 Example: Let $A = \{(a, .7), (b, .6), (c, .5), (d, [illegible])\}$ and $B = \{(a, .6), (b, [illegible]), (c, .2), (d, .9)\}$ be two fuzzy subsets of $X = \{a, b, c, d\}$.

Then $A - B = \{(a, .4), (b, [illegible]), (c, [illegible]), (d, .1)\}$.

[illegible] Defn: Let A and B be any two fuzzy subsets of a universal set X. Then the symmetric difference of A and B is defined as

$$A \Delta B = (A - B) \cup (B - A)$$

$$= (A \cap \bar{B}) \cup (\bar{B} \cap A)$$

$$= \max\{\min\{A(x), 1-B(x)\}, \min\{1-A(x), B(x)\}\}$$

[illegible] Example: Let $A = \{(a, .3), (b, .5), (c, .1)\}$ and $B = \{(a, .5), (b, .4), (c, .6)\}$ be two fuzzy subsets of $X = \{a, b, c\}$

$$A - B = \{(a, .3), (b, .5), (c, .1)\}$$

$$B - A = \{(a, .5), (b, .4), (c, .6)\}$$

$$\therefore A \Delta B = \{(a, .5), (b, .5), (c, .6)\}.$$

Note: ① The law of contradiction is violated for fuzzy sets. $A \cap \bar{A} = \phi$

i.e. $A \cap \bar{A} \neq \phi$ or $\min\{A(x), 1-A(x)\} = 0$

(since no elements satisfied the condition)

② The law of excluded middle is violated for fuzzy sets.

i.e. $A \cup \bar{A} = X$ or $\max\{A(x), 1-A(x)\} = 1$

(since no elements satisfied the condition)

problem:

U.Q 5m ① prove that Demorgan's law for the fuzzy sets.

proof: i) $(\overline{A \cup B}) = \bar{A} \cap \bar{B}$

ii) $(\overline{A \cap B}) = \bar{A} \cup \bar{B}$

i) Let $x \in X$.

NOW $\overline{A \cup B}(x) = 1 - (A \cup B)(x)$

$= 1 - \max\{A(x), B(x)\}$

$= \min \{1 - A(x), 1 - B(x)\}$

$= (\bar{A} \cap \bar{B})(x), \quad \forall x \in X$

$\therefore \overline{A \cup B} = \bar{A} \cap \bar{B}$

ii) Let $x \in X$

Now $\overline{A \cap B}(x) = 1 - (A \cap B)(x)$

$= 1 - \min \{A(x), B(x)\}$

$= \max \{1 - A(x), 1 - B(x)\}$

$= (\bar{A} \cup \bar{B})(x) \quad \forall x \in X.$

$\therefore \overline{A \cap B} = \bar{A} \cup \bar{B}$

1.49 **α-cut :** Let A be a fuzzy subset of a universal set X and $\alpha \in [0,1]$. Then α-cut is defined as $\alpha_A = \{x / A(x) \geq \alpha\}$. Clearly it is subset of X.

1.50 **Example :** Let $A = \{(a, .4), (b, .5), (c, .2), (d, .$ be a fuzzy subset of $X = \{a, b, c, d\}$. Then $.3_A = \{a, b, d\}$ and $.8_A = \{d\}$.

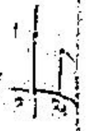
1.51 **Strong α-cut :** Let A be a fuzzy subset of universal set X. Then Strong α-cut is defined as $\alpha + A = \{x / A(x) > \alpha\}$. Clearly it is a subset of X.

1.52 **Example :** Let $A = \{(a, .5), (b, .2), (c, .3), (d,$ be a fuzzy subset of $X = \{a, b, c, d\}$. Thus $.4 + A = \{a, d\}$ and $.9 + A = \phi$.

153 Example: We consider three fuzzy subsets that represent the concepts of a young, middle aged and old person. The trapezoidal membership function A_1, A_2 and A_3 of the fuzzy subsets A_1, A_2, A_3 of the interval $[0, 80]$ are defined as

$$A_1(x) = \begin{cases} 1 & \text{when } x \le 20 \\ (35-x)/15 & \text{when } 20 < x < 35 \\ 0 & \text{when } x \ge 35 \end{cases}$$

$$A_2(x) = \begin{cases} 0 & \text{when either } x \le 20 \text{ or } \ge 60 \\ (x-20)/15 & \text{when } 20 < x < 35 \\ (60-x)/15 & \text{when } 45 < x < 60 \\ 1 & \text{when } 35 \le x \le 45 \end{cases}$$



and

$$A_3(x) = \begin{cases} 0 & \text{when } x \le 45 \\ (x-45)/15 & \text{when } 45 < x < 60 \\ 1 & \text{when } x \ge 60 \end{cases}$$

$${}^0A_1 = {}^0A_2 = {}^0A_3 = [0, 80]$$

$\alpha \in (0, 1]$, then

$${}^\alpha A_1 = [0, 35 - 15\alpha]$$

$${}^\alpha A_2 = [15\alpha + 20, 60 - 15\alpha]$$

$${}^\alpha A_3 = [15\alpha + 45, 80]$$

by $\frac{35-x}{15} \ge \alpha$

$35 - x \ge 15\alpha$

$\frac{x-20}{15} \ge \alpha$

$x - 20 \ge 15\alpha$

$x \ge 15\alpha + 20$

Similarly and

$${}^{\alpha+}A_1 = (0, 35 - 15\alpha)$$

$${}^{\alpha+}A_2 = (15\alpha + 20, 60 - 15\alpha)$$

$${}^{\alpha+}A_3 = (15\alpha + 45, 80).$$

Also $\quad {}^{1+}A_1 = {}^{1+}A_2 = {}^{1+}A_3 = \phi$

154 Level set of A:

Let A be a fuzzy subset of a universal set X. The level set of A is defined by as $\Lambda(A) = \{\alpha / A(x) = \alpha, \forall x \in X\}$

(or)

The set of all levels $\alpha \in [0,1]$ that represents distinct α-cuts of a fuzzy subset A is called a level set of A.

Example: let $A = \{(a, .4), (b, .6), (c, .8), (d, .2)\}$ be a fuzzy subset of $X = \{a, b, c, d, e\}$. Then $\Lambda_A = \{0, .2, .4, .6, .8\}$

Problem 2: Let A be a fuzzy subset of a set and $\alpha_1, \alpha_2 \in [0,1]$, $\alpha_1 < \alpha_2$. Show that

i) ${}^{\alpha_1}A \supseteq {}^{\alpha_2}A$ ii) ${}^{\alpha_1 +}A \supseteq {}^{\alpha_2 +}A$...

Proof: i) Let $x \in X$

Let $x \in {}^{\alpha_2}A$

$\Rightarrow A(x) \geq \alpha_2$

$\Rightarrow A(x) \geq \alpha_2 > \alpha_1$

$\Rightarrow A(x) \geq \alpha_1$

$\Rightarrow x \in {}^{\alpha_1}A$

$\therefore {}^{\alpha_2}A \subseteq {}^{\alpha_1}A$.

~~let $x \in X$~~

Let $x \in {}^{\alpha_2 +}A$

$\Rightarrow A(x) > \alpha_2$

$A(x) > \alpha_2 > \alpha_1$

$A(x) \geq \alpha_1$

$\Rightarrow x \in {}^{\alpha_1 +}A \Rightarrow {}^{\alpha_2 +}A \subseteq {}^{\alpha_1 +}A$

Result: Let A be a fuzzy subsets of X and $\alpha_1, \alpha_2 \in [0,1]$ such that $\alpha_1 < \alpha_2$, S.T

(i) ${}^{\alpha_1}A \cap {}^{\alpha_2}A = {}^{\alpha_2}A$

(ii) ${}^{\alpha_1}A \cup {}^{\alpha_2}A = {}^{\alpha_1}A$

iii) ${}^{\alpha_1 +}A \cap {}^{\alpha_2 +}A = {}^{\alpha_2 +}A$

iv) ${}^{\alpha_1 +}A \cup {}^{\alpha_2 +}A = {}^{\alpha_1 +}A$

Proof :

Given that $\alpha_1, \alpha_2 \in [0,1]$, $\alpha_1 < \alpha_2$

$\alpha_1 \subset \alpha_2$

$\Rightarrow \alpha_2 A \subset \alpha_1 A \rightarrow$ (1) and $\alpha_2 + A \subseteq \alpha_1 + A \rightarrow$ (2)

(i) $\alpha_1 A \cap \alpha_2 A = \alpha_2 A$ from (1)

(ii) $\alpha_1 A \cup \alpha_2 A = \alpha_1 A$

(iii) $\alpha_1 + A \cap \alpha_2 + A = \alpha_2 + A$

(iv) $\alpha_1 + A \cup \alpha_2 + A = \alpha_1 + A$

1.5.6 Q. convex fuzzy subset

A fuzzy subset A of a universal set X is called convex fuzzy subset if all α-cuts are convex where $\alpha \in (0,1]$.

(or)

A fuzzy subset A of a universal set X is called convex fuzzy subset if $A(\lambda x_1 + (1-\lambda) x_2) \geq \min \{A(x_1), A(x_2)\}$ and all $\lambda \in [0,1]$ $\forall x_1, x_2 \in X$

U.Q. 5m Thm (1) A fuzzy subset A on R is convex iff $A(\lambda x_1 + (1-\lambda) x_2) \geq \min \{A(x_1), A(x_2)\}$ for all $x_1, x_2 \in R$ and all $\lambda \in [0,1]$ where min denotes minimum operator.

Proof : Assume that A is convex and $\alpha = A(x_1) \leq A(x_2)$

Then $x_1, x_2 \in {}^{\alpha}A$ since ${}^{\alpha}A$ is convex.

$\Rightarrow \lambda x_1 + (1-\lambda) x_2 \in {}^{\alpha}A$ hence

$\Rightarrow A(\lambda x_1 + (1-\lambda) x_2) \geq \alpha = \min \{A(x_1), A(x_2)\}$

$\Rightarrow A(\lambda x_1 + (1-\lambda) x_2) \geq \min \{A(x_1), A(x_2)\}$

$\Rightarrow A(\lambda x_1 + (1-\lambda) x_2)$ and $\lambda \in [0,1]$

$\forall x_1, x_2 \in R$ and

Conversely, assume that
$$A(\lambda x_1 + (1-\lambda)x_2) \geq \min\{A(x_1), A(x_2)\}$$
$\forall x_1, x_2 \in R$ and $\lambda \in [0,1]$,

To P.T A is convex.

It is enough to prove, $^{\alpha}A$ is convex.

$\alpha \in (0,1]$.

Let $x_1, x_2 \in {}^{\alpha}A$.

$\Rightarrow A(x_1) \geq \alpha$ and $A(x_2) \geq \alpha$.

Now, $A(\lambda x_1 + (1-\lambda)x_2) \geq \min\{A(x_1), A(x_2)\}$ given.

$\geq \min\{\alpha, \alpha\}$

$= \alpha$.

$A(\lambda x_1 + (1-\lambda)x_2) \geq \alpha$.

$\Rightarrow \lambda x_1 + (1-\lambda)x_2 \in {}^{\alpha}A$

$\therefore {}^{\alpha}A$ is convex for all $\alpha \in [0,1]$.

Hence A is convex.

Note:

Any property generalized from classical set theory into the domain of fuzzy subset theory that is preserved in all α-cuts for $\alpha \in (0,1]$ in the classical sense is called a <u>cutworthy property</u>, if it is preserved in all strong α-cuts for $\alpha \in [0,1]$ in the classical sense is called a <u>strong cut worthy property</u>.

3) A fuzzy subset C of $X = \{0,1,2,\ldots,10\}$ is defined as $C(x) = \frac{x}{1+x}$ for $x \in X$. Find the scalar cardinality of C.

Soln: The scalar cardinality of C is

$$|C| = \sum C(x)$$
$$= \sum_{x=0}^{10} \frac{x}{1+x}$$
$$= 0 + \tfrac{1}{2} + \tfrac{2}{3} + \tfrac{3}{4} + \ldots + \tfrac{10}{11}$$
$$= 0 + 0.5 + 0.66 + 0.75 + 0.8 + 0.83 + 0.86 + 0.87 + 0.9 + 0.9$$
$$= 7.95$$

4) Let A, B be fuzzy subsets defined on a universal set X. P.T $|A|+|B| = |A \cup B| + |A \cap B|$ where $\cup, \cap$ mean union, intersection in fuzzy sense.

soln :-

Let $x \in X$

W.K.T $(A \cup B)(x) = Max\{A(x), B(x)\}$

$(A \cap B)(x) = Min\{A(x), B(x)\}$

$(A \cup B)(x) + (A \cap B)(x) = A(x) + B(x) \quad \forall x$

$\Rightarrow \sum (A \cup B)(x) + \sum (A \cap B)(x) = \sum A(x) + \sum B(x)$

$|A \cup B| + |A \cap B| = |A| + |B|$

5) A fuzzy subset A of $X = [0, 10]$ is defined as $A(x) = \frac{1}{1+x^2}$. Find the scalar cardinality of A.

soln :-

The Scalar Cardinality of A is

$$|A| = \int A(x) \cdot dx$$

$$= \int_0^{10} \frac{1}{1+x^2} \cdot dx$$

$$= \tan^{-1}(x) \Big]_0^{10}$$

$$= \tan^{-1}(10) - \tan^{-1}(0)$$

$$= \tan^{-1}(10) - 0$$

$$= 84.289.$$

6) S.T $(F(X), d)$ is a metric space with the Hamming distance $d(A, B)$ where $A, B \in F(X)$ (or) S.T the hamming distance $d(A, B)$ is a metric.

proof : Let $A, B, C \in F(X)$

W.K.T $d(A, B) = \sum_{x \in X} | A(x) - B(x) |$.

i) clearly $|A(x) - B(x)| \geq 0$

$$\Rightarrow \sum_{x \in X} |A(x) - B(x)| \geq 0$$

$\Rightarrow d(A,B) \geq 0 \quad \forall A,B \in F(X).$

ii) $d(A,B) = 0 \Leftrightarrow \sum_{x \in X} |A(x) - B(x)| = 0$

$\Leftrightarrow |A(x) - B(x)| = 0$

$\Leftrightarrow A(x) = B(x) \quad \forall x \in X.$

$\Leftrightarrow A = B$

$d(A,B) = 0 \Leftrightarrow A = B$

iii) $d(A,B) = \sum_{x \in X} |A(x) - B(x)|.$

$= \sum_{x \in X} |B(x) - A(x)|.$

$= d(B,A).$

$\therefore d(A,B) = d(B,A), \quad \forall A,B \in F(X).$

iv) Now, $|A(x) - C(x)| \leq |A(x) - B(x)| + |B(x) - C(x)|.$

$\Rightarrow \sum_{x \in X} |A(x) - C(x)| \leq \sum_{x \in X} |A(x) - B(x)| + \sum_{x \in X} |B(x) - C(x)|$

$\Rightarrow d(A,C) \leq d(A,B) + d(B,C), \quad \forall A,B,C \in F(X).$

$\therefore$ d is a metric.

Hence, $(F(X), d)$ is a metric space.

7) Consider the fuzzy sets A, B and C defined on the interval $X = [0,10]$ of real numbers by the membership grade functs $A(x) = \frac{x}{x+2}$, $B(x) = 2^{-x}$, $C(x) = \frac{1}{1+10(x-2)^2}$.

i) Determine mathematical formulas and graphs of the membership grade functs of each of the following sets.

a) $\overline{A}, \overline{B}, \overline{C}$ b) $A \cup B, B \cup C$ c) $A \cap B, B \cap C$, d) $A \cup B \cup C$, $A \cap B \cap C$

e) $A \cap \overline{C}$, $\overline{B \cap C}$, $\overline{A \cup C}$.

ii) Calculate the α-cuts for A and strong α-cuts for some values of α, $\alpha = 0.2, 0.5, 0.8$,

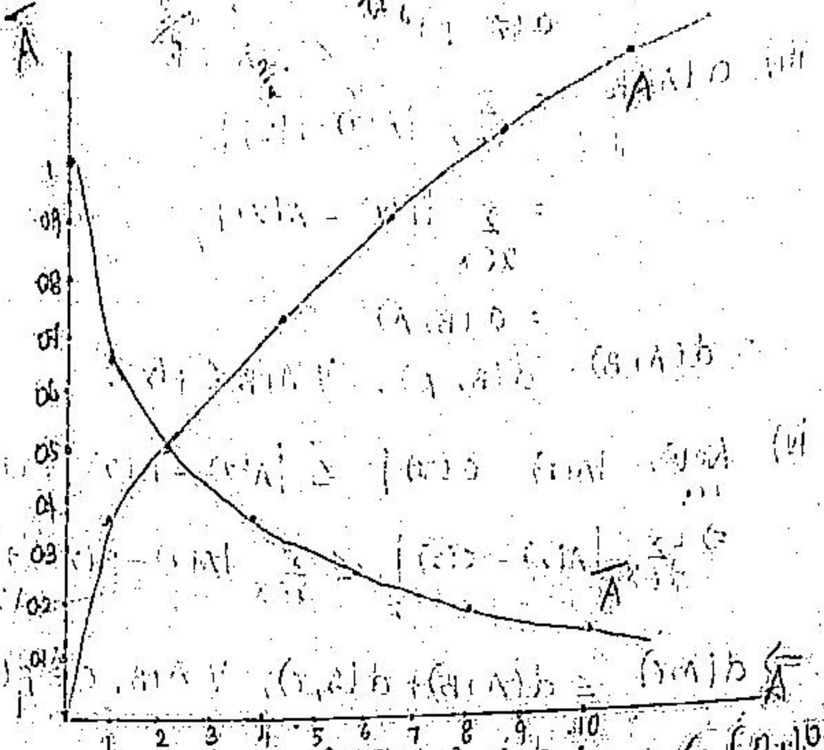
Soln:- Given that the membership value of the fuzzy set A is $A(x) = \frac{x}{x+2}$ $x \in [0,10]$.

$\therefore$ The membership value of the fuzzy set $\overline{A}$ is

$$\bar{A}(x) = 1 - A(x) = \frac{x+2-x}{x+2}$$

$$= 1 - \frac{x}{x+2} =$$

$$= \frac{2}{x+2}$$

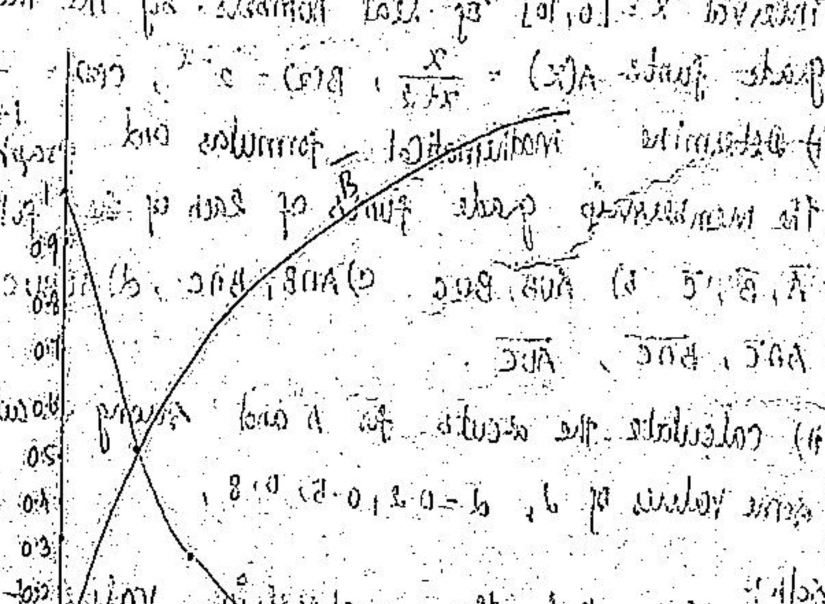


$x \in [0,10]$

$$B(x) = 2^{-x} = \frac{1}{2^x}$$

$$\bar{B}(x) = 1 - B(x) = 1 - \frac{1}{2^x} = \frac{2^x - 1}{2^x}$$

$x \in [0,10]$



$x \in [0,10]$

Given the membership value of the fuzzy set C is

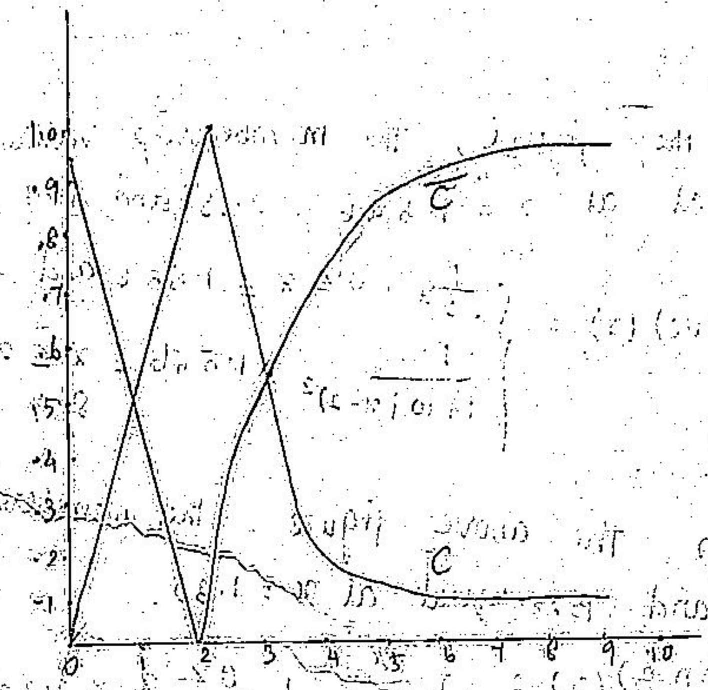
$$C(x) = \frac{1}{1+10(x-2)^2}, \quad x \in [0, 10]$$

∴ the membership value of the fuzzy set $\bar{C}$ is

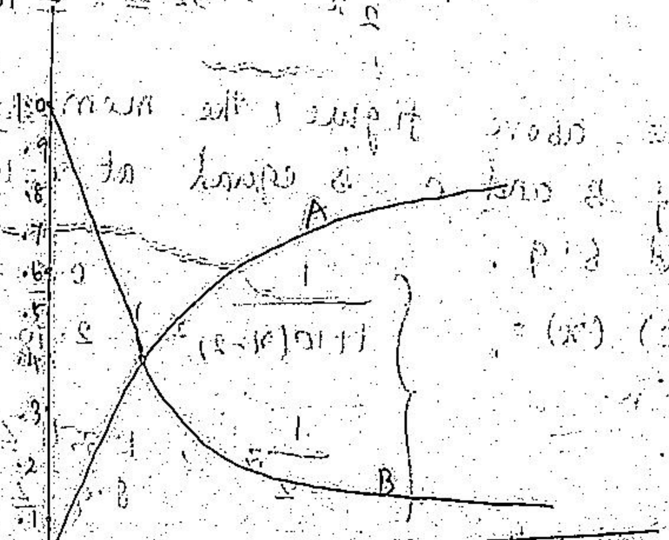
$$\bar{C}(x) = 1 - \frac{1}{1+10(x-2)^2}$$

$$= \frac{1+10(x-2)^2 - 1}{1+10(x-2)^2}$$

$$= \frac{10(x-2)^2}{1+10(x-2)^2}, \quad x \in [0, 10]$$

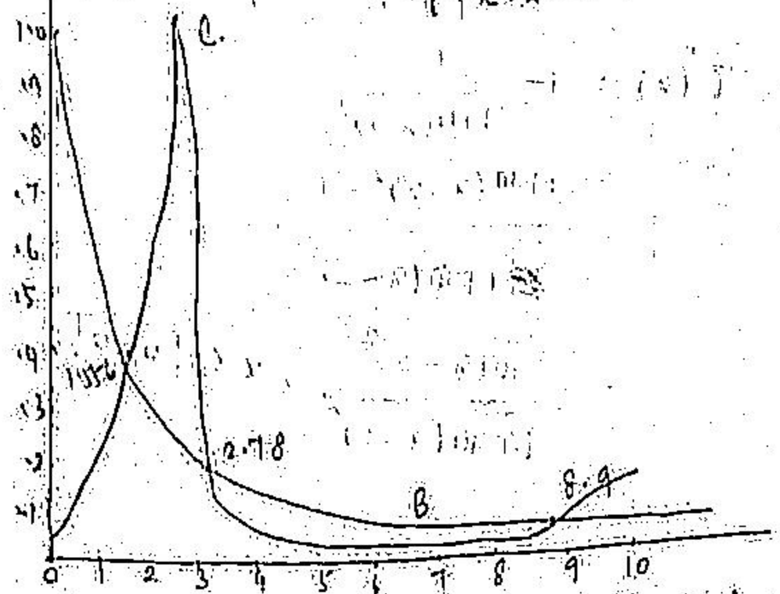


(b)



From the figure, the membership value of A and B is equal at $x = 1.32$ (using bisection Method)

$$\therefore (A \cup B)(x) = \begin{cases} \frac{1}{2^x}, & 0 \leq x \leq 1.32 \\ \frac{x}{x+2} & 1.32 \leq x \leq 10 \end{cases}$$



From the figure, the membership value of B and C is equal at $x = 1.556$, 2.78 and 8.9.

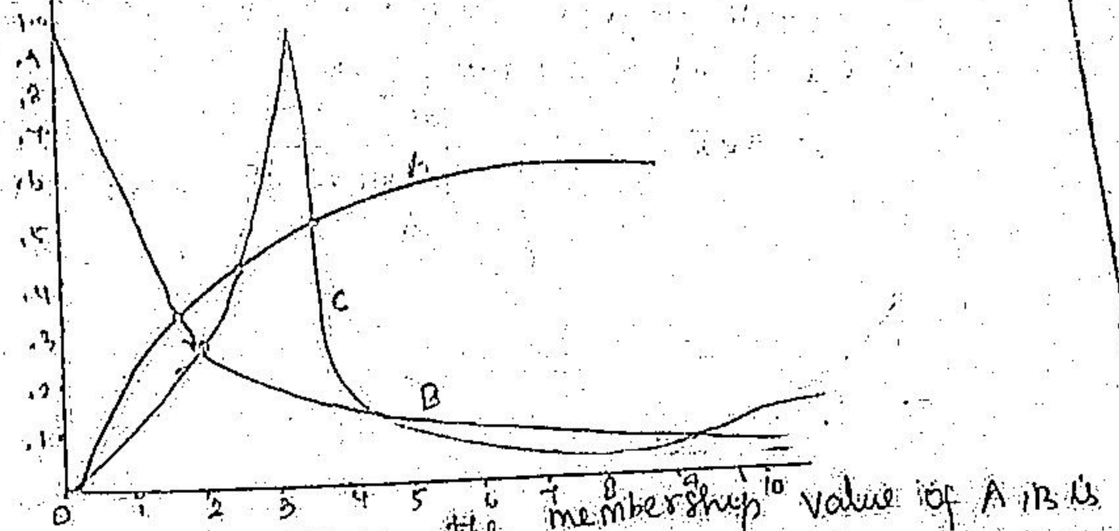
$$\therefore (B \cup C)(x) = \begin{cases} \frac{1}{2^x}, & 0 \leq x \leq 1.556 \text{ and } 2.78 \leq x \leq 8.9 \\ \frac{1}{1+10(x-2)^2}, & 1.556 \leq x \leq 2.78 \text{ and } 8.9 \leq x \leq 10 \end{cases}$$

ⓒ From the above figure the membership value of A and B is equal at $x = 1.32$.

$$\underset{\min}{(A \cap B)(x)} = \begin{cases} \frac{x}{x+2}, & 0 \leq x \leq 1.32 \\ \frac{1}{2^x}, & 1.32 \leq x \leq 10 \end{cases}$$

From the above figure, the membership value of B and C is equal at $x = 1.556$, 2.78 and 8.9.

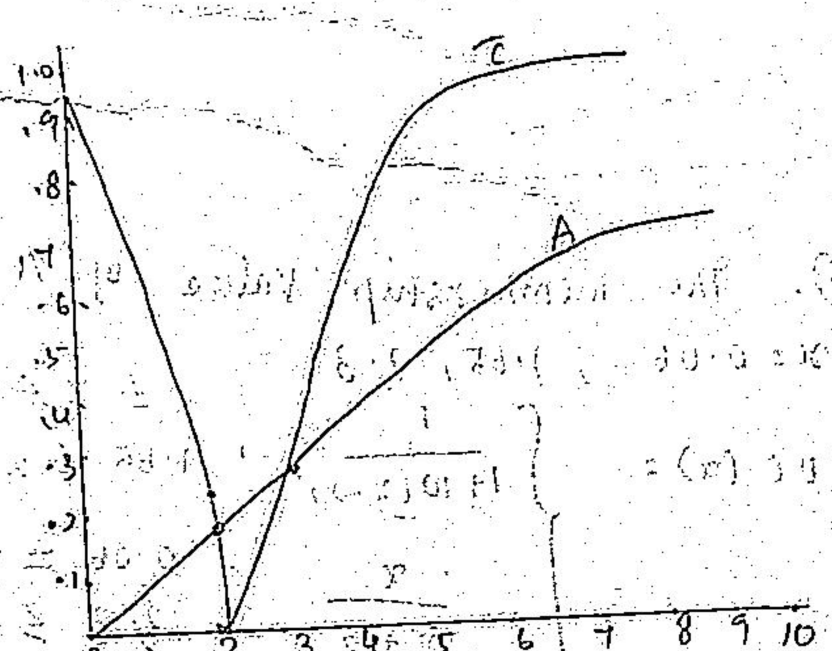
$$\underset{\text{Min}}{(B \cap C)(x)} = \begin{cases} \frac{1}{1+10(x-2)^2}, & 0 \leq x \leq 1.556 \text{ and } 2.78 \leq x \leq 8.9 \\ \frac{1}{2^x}, & 1.556 \leq x \leq 2.78 \text{ and } 8.9 \leq x \leq 10 \end{cases}$$



(a) From the figure, the membership value of A, B is equal at $x = 1.32$ and the membership value of B, C is equal at $x = 1.556, 2.78, 8.9$ and the membership value of A, C is equal at $x = 0.06, 1.65, 2.3$

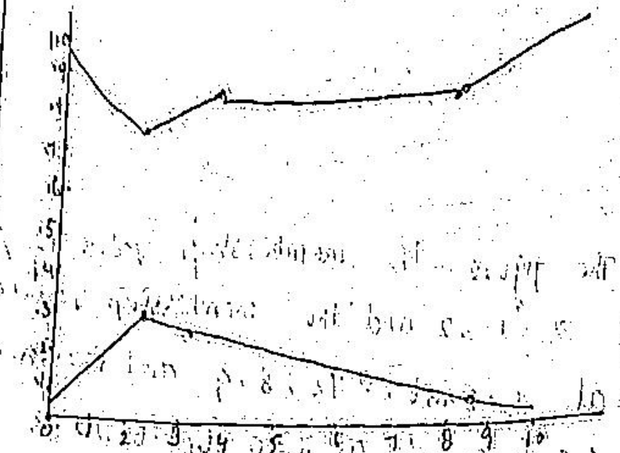
$$\therefore (A \cup B \cup C)(x) = \begin{cases} \frac{1}{2^x}, & 0 \le x \le 1.32 \\ \frac{x}{x+2} & 1.32 \le x \le 1.65 \text{ and } 2.3 \le x \le 10 \\ \frac{1}{1+10(x-2)^2}, & 1.65 \le x \le 2.3 \end{cases}$$

$$(A \cap B \cap C)(x) = \begin{cases} \frac{x}{x+2}, & 0 \le x \le 0.06 \\ \frac{1}{1+10(x-2)^2}, & 0.06 \le x \le 1.556 \text{ and } 2.78 \le x \le 8.9 \\ \frac{1}{2^x}, & 1.556 \le x \le 2.78 \text{ and } 8.9 \le x \le 10 \end{cases}$$



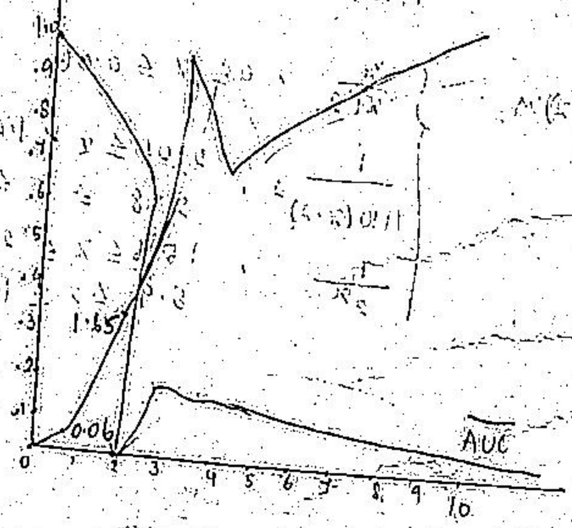
From the figure, the membership value of A and $\bar{C}$ is equal at $x = 1.108,\ 2.34$

$$\therefore A \cup \bar{C} = \begin{cases} \dfrac{10(x-2)^2}{1+10(x-2)^2}, & 0 \le x \le 1.108 \text{ and } 2.34 \le x \le 10 \\ \dfrac{x}{x+2}, & 1.108 \le x \le 2.34 \end{cases}$$



From (c), $\overline{B \cap C}(x) =$

$$\begin{cases} \dfrac{10(x-2)^2}{1+10(x-2)^2}, & 0 \le x \le 1.556 \text{ and } 2.78 \le x \le 8.9 \\ \dfrac{2^x - 1}{2^x}, & 1.556 \le x \le 2.78 \text{ and } 8.9 \le x \le 10 \end{cases}$$



From (d), the membership value of A and C is equal at $x = 0.06,\ 1.65,\ 2.3$

$$\therefore A \cup C(x) = \begin{cases} \dfrac{1}{1+10(x-2)^2}, & 0 \le x \le 0.06 \text{ and } 1.65 \le x \le 2.3 \\ \dfrac{x}{x+2}, & 0.06 \le x \le 1.65 \text{ and } 2.3 \le x \le 10 \end{cases}$$

$\therefore \lambda_{A \cup C}(x) = \begin{cases} \frac{10(x-2)^2}{11\,10(x-2)}, & 0 \le x \le 0.06 \text{ and } 1.65 \le x \le 2.3 \\ \frac{x}{x+2}, & 0.06 \le x \le 1.65 \text{ and } 2.3 \le x \le 10 \end{cases}$

(ii) To find α-cut in A :

If $\alpha = 0.2$, then $\frac{x}{x+2} = 0.2$

$\Rightarrow x = 0.2x + 0.4$

$\Rightarrow 0.8x = 0.4$

$\Rightarrow x = \frac{0.4}{0.8} = \frac{1}{2} = 0.5$

$\therefore {}^{0.2}A = \{x / 0.5 \le x \le 10\}$

If $\alpha = 0.5$ then $\frac{x}{x+2} = 0.5$

$\Rightarrow x = 0.5x + 1$

$\Rightarrow 0.5x = 1$

$\Rightarrow x = \frac{1}{0.5} = 2$

$\therefore {}^{0.5}A = \{x / 2 \le x \le 10\}$

If $\alpha = 0.8$, then $\frac{x}{x+2} = 0.8$

$\Rightarrow x = 0.8x + 1.6$

$\Rightarrow 0.2x = 1.6$

$x = \frac{16}{2} = 8$

$\therefore {}^{0.8}A = \{x / 8 \le x \le 10\}$

(iii) To find α-cut in B :

If $\alpha = 0.2$ then $\frac{1}{2^x} = 0.2$

$\Rightarrow \frac{1}{0.2} = 2^x$

$\Rightarrow 5 = 2^x$

$\Rightarrow \log 5 = x \log 2$

$x = \frac{\log 5}{\log 2} = 2.$

$\therefore {}_{2}B = \{x / 0 \le x \le 2\}$

If $\alpha = 0.5$, then $\frac{1}{2^x} = 0.5$

$\Rightarrow \frac{1}{0.5} = 2^x$

$\Rightarrow 2 = 2^x \Rightarrow x = 1$

$\Rightarrow {}_{0.5}B = \{x / 0 \le x \le 1\}$

If $\alpha = 0.8$, then $\frac{1}{2^x} = 0.8$

$\Rightarrow \frac{1}{0.8} = 2^x$

$\Rightarrow x = 0.32$

${}_{0.8}B = \{x / 0 \le x \le 0.32\}$

To find α cuts in C:

If $\alpha = 0.2$, then $\frac{1}{1+10(x-2)^2} = 0.2$

$\Rightarrow \frac{1}{0.2} = 1+10(x-2)^2$

$5-1 = 10(x-2)^2$

$\frac{4}{10} = (x-2)^2$

$0.4 = (x-2)^2$

$x-2 = \sqrt{0.4}$

$x-2 = \pm 0.6325$

$\Rightarrow x = 2.6325$ (or) 1.3675

$\therefore {}_{0.2}C = \{x / 1.3675 \le x \le 2.6325\}$

If $\alpha = 0.5$, then $\frac{1}{1+10(x-2)^2} = 0.5$

$\Rightarrow \frac{1}{0.5} = 1+10(x-2)^2$

$2 = 1+10(x-2)^2$

$\Rightarrow \frac{1}{10} = (x-2)^2$

$0.1 = (x-2)^2$

$x-2 = \pm 0.3162$

$\therefore x = 2.3162$ (or) 1.6838

$\therefore {}^{\cdot 5}C = \{x / 1.6838 \le x \le 2.3162\}$.

If $\alpha = 0.8$ then $\frac{1}{1+10(x-2)^2} = .8$

$\frac{1}{0.8} = 1+10(x-2)^2$

$1.25 - 1 = 10(x-2)^2$

$\frac{0.25}{10} = (x-2)^2$

$0.025 = (x-2)^2$

$x-2 = \pm 0.1581$

$x = 2.1581$ (or) 1.8419

$\therefore {}^{\cdot 8}C = \{x / 1.8419 \le x \le 2.158\}$

To find Strong α cut in A:

If $\alpha = 0.2$, then

${}^{0.2+}A = \{x / 0.5 < x \le 10\}$

If $\alpha = 0.5$, then

${}^{0.5+}A = \{x / 2 < x \le 10\}$

If $\alpha = 0.8$, then

${}^{0.8+}A = \{x / 8 < x \le 10\}$

To find Strong α cut in B:

If $\alpha = 0.2$, then ${}^{0.2+}B = \{x / 0 \le x < 2\}$

If $\alpha = 0.5$, then ${}^{0.5+}B = \{x / 0 \le x < 1\}$

If $\alpha = 0.8$, then ${}^{0.8+}B = \{x / 0 \le x < 0.32\}$

To find Strong α cut in C:

If $\alpha = 0.2$, then ${}^{0.2+}C = \{x / 1.3675 < x < 2.6325\}$

If $\alpha = 0.5$, then ${}^{0.5+}C = \{x / 1.6838 < x < 2.3162\}$

If $\alpha = 0.8$, then ${}^{0.8+}C = \{x / 1.8419 < x < 2.1581\}$

6) Calculate the degree of subsethood $S(C,D)$ and $S(D,C)$ for the fuzzy set C, D of $C(x) = \frac{x}{x+1}$, $D(x) = 1 - \frac{x}{10}$ for $x \in \{0,1,2,\ldots,10\} = X$, for $x = \{0,1,\ldots,10\}$

Soln: Given that $C(x) = \frac{x}{x+1}$, $x \in \{0,1,\ldots,10\}$

i.e. $C = \{(0,0), (1,0.5), (2,0.67), (3,0.75), (4,0.8), (5,0.83), (6,0.86), (7,0.88), (8,0.89), (9,0.9), (10,0.909)\}$.

Also given that

$D(x) = 1 - \frac{x}{10}$, $x \in \{0,1,\ldots,10\}$

$D = \{(0,1), (1,0.9), (2,0.8), (3,0.7), (4,0.6), (5,0.5), (6,0.4), (7,0.3), (8,0.2), (9,0.1), (10,0)\}$

$$|C| = \sum_{x=0}^{10} C(x) = 7.98$$

$$|D| = \sum_{x=0}^{10} D(x) = 5.5$$

$C \cap D = \{(0,0), (1,0.5), (2,0.67), (3,0.7), (4,0.6), (5,0.5), (6,0.4), (7,0.3), (8,0.2), (9,0.1), (10,0)\}$

$\therefore |C \cap D| = 3.97$

$$\therefore S(C,D) = \frac{|C \cap D|}{|C|} = \frac{3.97}{7.98} = 0.497$$

$$S(D,C) = \frac{|C \cap D|}{|D|} = \frac{3.97}{5.5} = 0.722$$

Theorem (2) Let $A, B \in F(X)$. Then the following properties hold for all $\alpha, \beta \in [0,1]$

(i) ${}^{\alpha+}A \subseteq {}^{\alpha}A$

(ii) $\alpha \leq \beta \Rightarrow {}^{\alpha}A \supseteq {}^{\beta}A$ and ${}^{\alpha+}A \supseteq {}^{\beta+}A$

(iii) ${}^{\alpha}(A \cap B) = {}^{\alpha}A \cap {}^{\alpha}B$ and ${}^{\alpha}(A \cup B) = {}^{\alpha}A \cup {}^{\alpha}B$

(iv) $\alpha+(A\cap B) = \alpha+A \cap \alpha+B$ and $\alpha+(A\cup B) = \alpha+A \cup \alpha+B$

v) $\alpha(\bar{A}) = (1-\alpha)+\bar{A}$.

proof:

i) Let $x \in X$ and $\alpha, \beta \in [0,1]$.

for all x, $x \in \alpha+A$

$\Rightarrow A(x) > \alpha$

$\Rightarrow A(x) \geq \alpha$

$\Rightarrow x \in {}^{\alpha}A$

$\therefore \alpha+A \subseteq {}^{\alpha}A$

ii) Given $\alpha \leq \beta$

for all x, $x \in {}^{\beta}A$

$\Rightarrow A(x) \geq \beta$

$\Rightarrow A(x) \geq \beta \geq \alpha$

$\Rightarrow A(x) \geq \alpha$

$\Rightarrow x \in {}^{\alpha}A$

$\therefore {}^{\beta}A \subseteq {}^{\alpha}A$

For all x, $x \in \beta+A$

$\Rightarrow A(x) > \beta$

$\Rightarrow A(x) \geq \beta \geq \alpha$

$\Rightarrow A(x) > \alpha$

$\Rightarrow x \in \alpha+A$

$\therefore \beta+A \subseteq \alpha+A$

iii) For any $x \in {}^{\alpha}(A\cap B)$

$\Leftrightarrow (A\cap B)(x) \geq \alpha$

$\Leftrightarrow \min\{A(x), B(x)\} \geq \alpha$

$\Leftrightarrow A(x) \geq \alpha$ and $B(x) \geq \alpha$.

$\Leftrightarrow x \in {}^{\alpha}A$ and $x \in {}^{\alpha}B$

$\Leftrightarrow x \in {}^{\alpha}A \cap {}^{\alpha}B$

$\therefore {}^{\alpha}(A\cap B) = {}^{\alpha}A \cap {}^{\alpha}B$

Note : $\alpha(A) \neq \alpha^{+}A$

Theorem (?) Let $A_i \in I(X)$ for all $i \in I$, where I is an index set. Then (i) $\bigcup_{i \in I} {}^{\alpha}A_i \subseteq {}^{\alpha}(\bigcup_{i \in I} A_i)$

and $\bigcap_{i \in I} {}^{\alpha}A_i = {}^{\alpha}(\bigcap_{i \in I} A_i)$.

(ii) $\bigcup_{i \in I} {}^{\alpha+}A_i = {}^{\alpha+}(\bigcup_{i \in I} A_i)$ and $\bigcap_{i \in I} {}^{\alpha+}A_i \supseteq {}^{\alpha+}(\bigcap_{i \in I} A_i)$.

Proof: Let $x \in X$.

i) For any $x \in \bigcup_{i \in I} {}^{\alpha}A_i$

$\Rightarrow \exists\, i_0 \in I$ s.t. $x \in {}^{\alpha}A_{i_0}$

$\Rightarrow A_{i_0}(x) \geq \alpha$

$\Rightarrow \sup_{i \in I} A_i(x) \geq \alpha$

$\Rightarrow \bigcup_{i \in I} A_i(x) \geq \alpha$.

$\Rightarrow x \in {}^{\alpha}(\bigcup_{i \in I} A_i)$

For any $x \in \bigcap_{i \in I} {}^{\alpha}A_i$

$\Leftrightarrow x \in {}^{\alpha}A_i,\ \forall\, i$

$\Leftrightarrow A_i(x) \geq \alpha,\ \forall\, i$

$\Leftrightarrow \inf A_i(x) \geq \alpha$

$\Leftrightarrow \bigcap_{i \in I} (x) \geq \alpha$

$\Leftrightarrow x \in {}^{\alpha}(\bigcap_{i \in I} A_i)$

$\bigcap_{i \in I} {}^{\alpha}A_i = {}^{\alpha}(\bigcap_{i \in I} A_i)$.

ii) For any $x \in \bigcup_{i \in I} {}^{\alpha+}A_i$

$\Leftrightarrow \exists\, i_0 \in I$ s.t. $x \in {}^{\alpha+}A_{i_0}$

$\Leftrightarrow A_{i_0}(x) > \alpha$.

$\Leftrightarrow \sup_{i \in I} A_i(x) > \alpha$

$\Leftrightarrow \left(\bigcup_{i \in I} A_i\right)(x) > \alpha$

$\Leftrightarrow x \in {}^{\alpha+}\left(\bigcup_{i \in I} A_i\right)$

$\therefore \bigcup_{i \in I} {}^{\alpha+}A_i = {}^{\alpha+}\left(\bigcup_{i \in I} A_i\right)$

for any $x \in {}^{\alpha+}\left(\bigcap_{i \in I} A_i\right)$

$\Rightarrow \left(\bigcap_{i \in I} A_i\right)(x) > \alpha$

$\Rightarrow \inf_{i \in I} A_i(x) > \alpha$

$\Rightarrow A_i(x) > \alpha \;\forall i$

$\Rightarrow x \in {}^{\alpha+}A_i, \;\forall i$

$\Rightarrow x \in \bigcap_{i \in I} {}^{\alpha+}A_i$

$\therefore {}^{\alpha+}\left(\bigcap_{i \in I} A_i\right) \subseteq \bigcap_{i \in I} {}^{\alpha+}A_i$.

1.57 Example ①

Let X be a universal set and let $A_i \in F(X)$ be defined by $A_i(x) = 1 - \frac{1}{i}$ for all $x \in X$ and all $i \in N$. Verify $\bigcup_{i \in N} {}^{\alpha}A_i \subseteq {}^{\alpha}\left(\bigcup_{i \in N} A_i\right)$.

Soln: Given that $A_i(x) = 1 - \frac{1}{i}$, $\forall x$ and all $i \in N$.

Then for any $x \in X$,

$\left(\bigcup_{i \in N} A_i\right)(x) = \sup_{i \in N} A_i(x) = \sup_{i \in N}\left(1 - \frac{1}{i}\right)$

$= 1$

Let $\alpha = 1$, ${}^{1}\left(\bigcup_{i \in N} A_i\right) = X$

For any i, ${}^{1}A_i = \phi$ because,

$A_i(x) = 1 - \frac{1}{i} < 1$

$\therefore \bigcup_{i \in N}^{\alpha} A_i = \phi \subset X = {}^{\alpha}\left(\bigcup_{i \in N} A_i\right)$.

In general,

$$\bigcup_{i \in N} {}^{\alpha}A_i \subseteq {}^{\alpha}\left(\bigcup_{i \in N} A_i\right)$$

1.58 Example ⓐ: Let X be a universal set and let $A_i \in F(X)$ be defined by $A_i(x) = \frac{1}{i}$ for all $x \in X$ and all $i \in N$. Verify ${}^{\alpha+}\left(\bigcap_{i \in N} A_i\right) \subset \bigcap_{i \in N} {}^{\alpha+}A_i$.

Soln: Given that $A_i(x) = \frac{1}{i}$ $\forall x$ and $i \in N$.

Then for any $x \in X$,

$$\left(\bigcap_{i \in N} A_i\right)(x) = \inf_{i \in N} A_i(x)$$

$$= \inf_{i \in N} \left(\frac{1}{i}\right)$$

$$= 0$$

Let $\alpha = 0$, ${}^{0+}\left(\bigcap_{i \in N} A_i\right) = \phi$

$\bigcap {}^{0+}A_i = X$ because,

for any i, $A_i(x) = \frac{1}{i} > 0$

${}^{0+}\left(\bigcap_{i \in N} A_i\right) = \phi \subseteq X = \bigcap_{i \in N} {}^{0+}A_i$

In general,

$$\bigcap_{i \in N} {}^{\alpha}A_i \subseteq {}^{\alpha}\left(\bigcap_{i \in N} A_i\right)$$

Theorem ④: Let $A, B \in F(X)$. Then for all $\alpha \in [0,1]$

U.Q 10M

i) $A \subseteq B$ iff ${}^{\alpha}A \subseteq {}^{\alpha}B$

ii) $A \subseteq B$ iff ${}^{\alpha+}A \subseteq {}^{\alpha+}B$

iii) $A = B$ iff ${}^{\alpha}A = {}^{\alpha}B$.

iv) $A = B$ iff ${}^{\alpha+}A = {}^{\alpha+}B$.

proof: $A, B \in F(X)$

(i) Given that A, B

Assume that $A \subseteq B$

i.e., $A(x) \le B(x) \ \forall x$

To prove that $\alpha_A \subseteq \alpha_B$.

Suppose if $\exists\, \alpha_0 \in [0,1]$ s.t. $\alpha_{0A} \not\subseteq \alpha_{0B}$

$\Rightarrow \exists\, x_0 \in X$ s.t. $x_0 \in \alpha_{0A}$ and $x_0 \notin \alpha_{0B}$

$\Rightarrow A(x_0) \ge \alpha_0$ and $B(x_0) < \alpha_0$

$\Rightarrow B(x_0) < \alpha_0 \le A(x_0)$

$\Rightarrow B(x_0) < A(x_0)$ which contradicts that $A \subseteq B$.

$\therefore \alpha_A \subseteq \alpha_B$

(or)

Let $x \in \alpha_A$

$\Rightarrow A(x) \ge \alpha$

$\Rightarrow B(x) \ge A(x) \ge \alpha$.

$\Rightarrow B(x) \ge \alpha$

$\Rightarrow x \in \alpha_B$

Hence $\alpha_A \subseteq \alpha_B$.

Conversely assume that $\alpha_A \subseteq \alpha_B$

i.e., for any $x \in X$, $\forall \alpha$, $x \in \alpha_A \subseteq \alpha_B$

for any $x \in X$, $\forall \alpha$, $A(x) \ge \alpha$ and $B(x) \ge \alpha$.

To prove that $A \subseteq B$.

For any $x \in X$, $A(x) = \alpha$

$\Rightarrow x \in \alpha_A$

$\Rightarrow x \in \alpha_B$

$\Rightarrow B(x) \ge \alpha$

$\Rightarrow B(x) \ge A(x) \ \forall x$

$\therefore A \subseteq B$

(or)

Assume that $\alpha_A \subseteq \alpha_B$

To prove that $A \subseteq B$.

suppose $A \nsubseteq B$.

$\Rightarrow \exists\, x_0 \in X$ s.t $A(x_0) > B(x_0)$

let $\alpha = A(x_0) \Rightarrow x_0 \in \alpha_A$.

$\Rightarrow \alpha > B(x) \Rightarrow x_0 \notin \alpha_B$

$\Rightarrow \alpha_A \nsubseteq \alpha_B$ which contradicts $\alpha_A \subseteq \alpha_B$.

$\therefore A \subseteq B$.

$\therefore A \subseteq B$ iff $\alpha_A \subseteq \alpha_B$.

(ii) Assume that $A \subseteq B$.

To prove that ${}^{\alpha+}A \subseteq {}^{\alpha+}B$

Let $x \in {}^{\alpha+}A$

$\Rightarrow A(x) > \alpha$

$\Rightarrow B(x) \geq A(x) > \alpha$

$\Rightarrow B(x) > \alpha$

$\Rightarrow x \in {}^{\alpha+}B$.

$\Rightarrow {}^{\alpha+}A \subseteq {}^{\alpha+}B$.

Conversely assume that ${}^{\alpha+}A \subseteq {}^{\alpha+}B$.

To prove that $A \subseteq B$.

suppose $A \nsubseteq B$.

$\Rightarrow \exists\, x_0 \in X$ s.t $A(x_0) > B(x_0)$

Let α be lies between $A(x_0)$ and $B(x_0)$

$\Rightarrow x_0 \in {}^{\alpha+}A$ and $x_0 \notin {}^{\alpha+}B$

$\Rightarrow {}^{\alpha+}A \nsubseteq {}^{\alpha+}B$, which contradicts ${}^{\alpha+}A \subseteq {}^{\alpha+}B$

$A \subseteq B$.

Hence $A \subseteq B$ iff ${}^{\alpha+}A \subseteq {}^{\alpha+}B$

(iii) Assume that $A = B$.

i.e) $A \subseteq B$ and $B \subseteq A$.

To prove that $\alpha_A = \alpha_B$

It is enough to prove that $\alpha_A \subseteq \alpha_B$ and α_B

If $A \subseteq B$, then $\alpha_A \subseteq \alpha_B$ by (ii)

If $B \subseteq A$, then $\alpha_B \subseteq \alpha_A$ by (ii)

$\therefore \alpha_A = \alpha_B$

Conversely assume that $\alpha_A = \alpha_B$

i.e. $\alpha_A \subseteq \alpha_B$ and $\alpha_B \subseteq \alpha_A$

To prove that $A = B$

It is enough to prove that

$A \subseteq B$ and $B \subseteq A$

If $\alpha_A \subseteq \alpha_B$, then $A \subseteq B$ by (ii)

If $\alpha_B \subseteq \alpha_A$, then $B \subseteq A$ by (ii)

$\therefore A = B$

Hence $A = B$ iff $\alpha_A = \alpha_B$

iv) Assume that $A = B$

i.e. $A \subseteq B$ and $B \subseteq A$

To prove that $\alpha + A = \alpha + B$

It is enough to prove that

$\alpha + A \subseteq \alpha + B$ and $\alpha + B \subseteq \alpha + A$

If $A \subseteq B$, then $\alpha + A \subseteq \alpha + B$ by (iii)

If $B \subseteq A$, then $\alpha + B \subseteq \alpha + A$ by (iii)

$\therefore \alpha + A = \alpha + B$

Conversely assume that $\alpha + A = \alpha + B$

i.e. $\alpha + A \subseteq \alpha + B$ and $\alpha + B \subseteq \alpha + A$

To prove that $A = B$

It is enough to prove that $A \subseteq B$ and $B \subseteq A$

If $\alpha + A \subseteq \alpha + B$, thus $A \subseteq B$ by (iii)

If $\alpha + B \subseteq \alpha + A$, thus $B \subseteq A$ by (iii)

$\therefore A = B$

Hence $A = B$ iff $\alpha + A = \alpha + B$

$= \bigcup_{\alpha \in [0,1]} \alpha \cdot [f(A) ...$

$= \bigcup_{\alpha \in [0,1]} \alpha \cdot [f({}^{\alpha+}A)]$, by above thm (i)

$= \bigcup_{\alpha \in [0,1]} f({}_{\alpha+}A)$.

UNIT - II

2m ✓ Fuzzy complement :

Let A be fuzzy set on X. Then the fuzzy complement CA, C_A or C of A is defined as $C : [0,1] \to [0,1]$ which ~~gi~~ assigns a values $C(A(x))$ to each membership grade $A(x)$ of any given fuzzy set A that is $c(\ldots)$ [illegible] for all $a \in X$.

The fuzzy complement c satisfy the following axioms

(i) $c(0)=1$, $c(1)=0$ (boundary condition)

(ii) for all $a, b \in [0,1]$, if $a \leq b$, then $c(a) \geq c(b)$ (monotonicity)

(iii) c is a continuous function

(iv) c is involutive, which means that $c(c(a)) = a$ $\forall\, a \in [0,1]$

Remark ① To produces meaningfull fuzzy complement's function c must satisfy atleast first two axioms.

② Axioms c_1 and c_2 are called the axiomatic skeleton for fuzzy complements

Theorem 2.1: Let a function $c : [0,1] \to [0,1]$ satisfy the axioms (ii) and (iv) thus c also satisfies axioms (i) and (iii). Moreover c must be a bijective function.

Proof: since the range of c is $[0,1]$,

$c(0) \leq 1$ and $c(1) \geq 0$ → ①

$\Rightarrow c(0) \leq 1$

$\Rightarrow c(c(0)) \geq c(1)$, by axiom (ii)

$\Rightarrow 0 \geq c(1)$, by axiom (iv) → ②

$c(1) = 0$

Consider $c(1)$

$\Rightarrow c(c(1)) = c(0)$

$\Rightarrow 1 = c(0)$

$\therefore c(0) = 1$ and $c(1) = 0$

Axiom (i) is proved.

Next we prove that c is bijective function

for every $a \in [0,1]$, $\exists$ $b \in [0,1]$ s.t

$$c(a) = b$$

$$\Rightarrow c(c(a)) = c(b)$$

$$\Rightarrow a = c(b)$$

$\therefore$ c is onto.

Suppose $c(a_1) = c(a_2)$

$$\Rightarrow c(c(a_1)) = c(c(a_2))$$

$$\Rightarrow a_1 = a_2$$

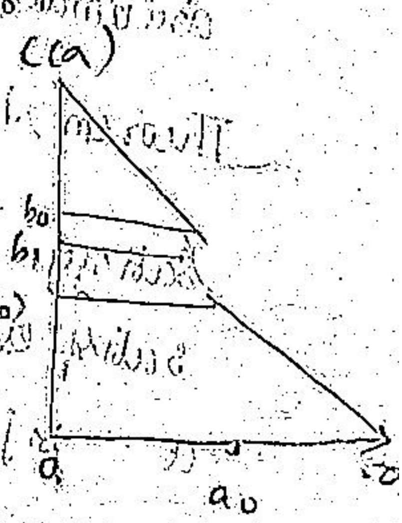
$\therefore$ c is 1-1

Hence c is bijective function.

Next we prove that c is continuous.

Since c is bijective and if $a \le b$ then $c(a) \geq c(b)$. Clearly c is continuous.

Suppose c is discontinuity at a_0



From the diagram

$$b_0 = \lim_{a \to a_0} c(a) > c(a_0)$$

Clearly there must exist $b_1 \in [0,1]$ such that $b_0 > b_1 > c(a_0)$

for which no $a_1 \in [0,1]$ exist s.t $c(a_1) = b_1$

This contradicts to the fact of the bijective function.

Hence c is continuous.

Note:

for $w=1$, the function becomes the classical fuzzy complement.

Defn:

Let c be a fuzzy complement of A. Then the equilibrium of a fuzzy complement c is defined as any value a for which $c(a)=a$.

Note:

The equilibrium value of the classical fuzzy complement is 0.5.

Theorem: 2.2

Every fuzzy complement has atmost one equilibrium.

Proof:

Let c be a arbitrary fuzzy complement. An equilibrium of c is a solution of the equation $c(a)-a=0$

Suppose a_1, a_2 s.t $a_1 < a_2$ are equilibrium of c.

$c(a_1)-a_1=0$ and

$c(a_2)-a_2=0$

$\therefore c(a_1)-a_1 = c(a_2)-a_2 \rightarrow (1)$

But $a_1 < a_2$

$\Rightarrow c(a_1) > c(a_2)$

$\Rightarrow c(a_1)-a_1 > c(a_2)-a_1$

$\Rightarrow c(a_1)-a_1 > c(a_2)-a_1 > c(a_2)-a_2$

$\Rightarrow c(a_1)-a_1 > c(a_2)-a_2$

Which $\Rightarrow \Leftarrow$ to (1).

Hence every fuzzy complement has atmost one equilibrium

Thm 2.3:

Assume that the gn fuzzy complement c has a equilibrium e_c, which is unique.

then (i) $a \le c(a)$ iff $a \le e_c$

(ii) $a \ge c(a)$ iff $a \ge e_c$

proof: Given that the fuzzy complement c has an equilibrium e_c, which is unique.

i.e. $c(e_c) = e_c \rightarrow (1)$

i) Assume that $a \le e_c$

$\Rightarrow c(a) \ge c(e_c)$

$= e_c$ by (1)

$\ge a$

$\Rightarrow c(a) \ge a$

i.e. $a \le c(a)$

conversely assume that $a \le c(a)$ To p.T $a \le e_c$

suppose $e_c < a$

$\Rightarrow c(e_c) > c(a)$

$\Rightarrow e_c > c(a)$ by (1)

$\ge a$ by assumption

$\Rightarrow e_c > a$

which $\Rightarrow \Leftarrow$ to $e_c < a$.

$a \le e_c$

$a \le e_c$ iff $a \le c(a)$

ii) Assume that $a \ge e_c$

$\Rightarrow c(a) \le c(e_c)$

$= e_c$ by (1)

$\le a$

$\Rightarrow c(a) \le a$

$a \ge c(a)$

conversely assume that $a \ge c(a)$ to p.T $a \ge e_c$.

suppose $e_c > a$

$\Rightarrow c(e_c) < c(a)$

$\Rightarrow e_c < c(a)$ by (1)

$\leq a$ by assumption

$\Rightarrow e_c \leq a$ to $e_c > a$

which $\Rightarrow \Leftarrow$ to $e_c > a$

$a \geq e_c$ iff $a \geq e_c$.

Hence $a \geq c(a)$ iff

Thm 2.4: If c is a continuous fuzzy complement then c has a unique equilibrium.

proof: The equilibrium e_c of a fuzzy complement c is the solution of the equation.

$$c(a) - a = 0$$

This is the special case of more general form of the eqn

$$c(a) - a = b, \text{ where } b \in [-1, 1].$$

since $c(1) - 1 = 0 - 1 = -1$ and

$c(0) - 0 = 1 - 0 = 1$

Since c is a continuous complement, it follows from the intermediate value thm for continuous function that for each $b \in [-1, 1]$, $\exists$ atleast one a such that $c(a) - a = b$

This demonstrates the necessary existence of an equilibrium value for a continuous function and thm 2.2 guarantees its uniqueness.

Defn: Let c be fuzzy complement. A membership grade of an element is represented by a real number $a \in [0, 1]$. Then any membership grade represented by the real number $d_a \in [0, 1]$ s.t $c(d_a) - d_a = a - c(a)$, is called a dual point of a with respect to c.

Theorem 2.5: If a complement c has an equilibrium e_c, then $d_{e_c} = e_c$.

proof

If $a = e_c$ thus by our definition of equilibrium $c(a) = a \Rightarrow a - c(a) = 0 \rightarrow (1)$

If $d_a = e_c$ then by our definition of equilibrium

$$c(d_a) = d_a$$

$$\Rightarrow c(d_a) - d_a = 0 \rightarrow (2)$$

From (1) and (2) we get

$$c(d_a) - d_a = a - c(a)$$

$\therefore$ d_a is a dual point of a when

$$a = e_c = d_a = de_c$$

$$de_c = e_c$$

Theorem For each $a \in [0,1]$ $d_a = c(a)$ iff $c(c(a)) = a$, that is where the complex is involutive.

proof Let $d_a = c(a)$, d_a is a dual point of a.

W.K.T if d_a is a dual point of a

then $c(d_a) - d_a = a - c(a) \rightarrow (1)$

put $d_a = c(a)$

$(1) \Rightarrow c(c(a)) - c(a) = a - c(a)$

$\Rightarrow c(c(a)) = a$

conversely assume that $c(c(a)) = a$

put $a = c(c(a))$ in (1) we get

$c(d_a) - d_a = c(c(a)) - c(c(c(a)))$

$\Rightarrow = c(c(a)) - c(a) \rightarrow (2)$

$d(a) = c(a)$ is a solution of (2)

$$\therefore d_a = c(a)$$

(8) Theorem 2.7: (First characterization Theorem of fuzzy complements).

Statement: Let c be a function from $[0,1]$ to $[0,1]$. Then c is a fuzzy complement (involutive) iff there exists a continuous function g from $[0,1]$ to R such that $g(0)=0$, g is strictly increasing and $c(a) = g^{-1}(g(1) - g(a))$ for all $a \in [0,1]$.

Proof:

Assume that g is continuous function from $[0,1]$ to R s.t $g(0)=0$. g is strictly increasing and

$$c(a) = g^{-1}(g(1) - g(a)) \ \forall a \in [0,1].$$

To prove that c is a fuzzy complement.

Then $g^{(-1)} : R \to [0,1]$ is defined by

$$g^{(-1)}(a) = \begin{cases} 0 & \text{if } a \in (-\infty, 0) \\ g^{-1}(a) & \text{if } a \in [0, g(1)] \\ 1 & \text{if } a \in (g(1), \infty) \end{cases}$$

When g^{-1} is an ordinary inverse of g.

Let c be a function on $[0,1]$ defined by $c(a) = g^{-1}(g(1) - g(a))$, $\forall a$. First we show that c satisfies axiom $c(ii)$.

Let c be a function on $[0,1]$ defined a

$c(a) = g^{-1}(g(1) - g(a))$, $\forall a$

First we show that c satisfies axiom c_2

For any $a, b \in [0,1]$, if $a < b$.

$\Rightarrow g(a) < g(b)$, g is strictly increasing

$\Rightarrow g(1) - g(a) > g(1) - g(b)$

$\Rightarrow g^{-1}(g(1) - g(a)) \geq g^{-1}(g(1) - g(b))$

$\Rightarrow c(a) > c(b)$

If $a < b$, then $c(a) > c(b)$

$\therefore$ the axiom (ii) is satisfied.

Next we show that c is involutive.

(iii) $c(c(a)) = a$

For any $a \in [0,1]$, then.

$c(c(a)) = g^{-1}[g(1) - g(c(a))]$

$= g^{-1}[g(1) - g[g^{-1}[g(1) - g(a)]]]$

$= g^{-1}[g(1) - g(1) + g(a)]$

$= g^{-1}[\cancel{g(1)} - \cancel{g(1)} + g(a)]$

$= a$

$\therefore c$ satisfies the axiom (iv).

By theorem 2.1, c also satisfies the axioms (i) and (iii). c is fuzzy complement. Conversely, c is a fuzzy complement.

assume that

To prove that $g: [0,1] \to R$

(i) g is continuous

(ii) g is strictly increasing.

(iii) $c(a) = g^{-1}(g(1) - g(a)) \; \forall a \in [0,1]$.

by theorem 2.4, c must have a unique equilibrium, let us say e_c.

$c(e_c) = e_c$, $e_c \in (0,1)$.

let $h: [0, e_c] \to [0, b]$ be any continuous strictly increasing bijection such that $h(0) = 0$ and $h(e_c) = \dots$ where b is any fixed positive real numbers (eg $h(a) \cdot \frac{ab}{ec}$)

Now we define a function $g : [0,1] \to [0,1]$

$$g(a) = \begin{cases} h(a) & \text{if } a \in [0, e_c] \\ 2b - h(c(a)) & a \in [e_c, 1] \end{cases}$$

(i) clearly $g(a) = 0$. since $0 \in [0, e_c]$ so $g(0) = h(0) = 0$

(ii) clearly g is continuous since h is continuous

(iii) clearly g is strictly increasing. since h is strictly increasing.

It is easy to s.t the pseudo inverse of g is given by

$$g^{(-1)}a = \begin{cases} a & \text{if } a \in (-\infty, 0) \\ h^{-1}(a) & \text{if } a \in [0, b] \\ c(h^{-1}(2b-a)) & \text{if } a \in [b, 2b] \\ 1 & \text{if } a \in (2b, \infty) \end{cases} \rightarrow ②$$

For any $a \in [0,1]$ if $a \in [0, e_c]$ then

$g^{-1}(g(1) - g(a)) = g^{-1}(2b - g(a))$

$= [\because g(1) = 2b - h(c(a))$ by ①

$= [g^{-1}(2b - h(a))$ by ①

$= 2b - h(a)$

$= c(h^{-1}(2b - (2b - h(a))) \quad \boxed{= 2b - 0 = 2b}$

$= c(h^{-1}(h(a)))$

$= c(a)$

$C(a) = g^{-1}(g(1) - g(a))$

And any $a \in [0,1]$ if $a \in [e_c, 1]$ then

$g^{-1}(g(1) - g(a)) = g^{-1}(2b - g(a)) = [\because g(1) = 2b]$

$= g^{-1}(g(1) - (g(1) - h(c(a)))$ by (1)

$= g^{-1}(h(c(a)))$

$= h^{-1}(h(c(a)))$ by (2)

$= c(a)$

$\therefore c(a) = g^{-1}(g(1) - g(a))$

Def:

A function $g : [0,1] \to R$ is defined by (i) $g(0) = 0$ (ii) g is continuous (iii) g is strictly increasing and (iv) $c(a) = g^{-1}(g(1) - g(a))$. The function g is called increasing generator.

Note: (1) The increasing generators of sugeno class of fuzzy complements are

$$g(a) = \frac{1}{\lambda} \ln(1+\lambda a) \text{ for } \lambda > -1.$$

(2) The increasing generator of the yager class of fuzzy complements are $g_w(a) = a^w$ for $w > 0$.

(3) The class of two parameters increasing generators $g_{\lambda w}(a) = \frac{1}{\lambda} \ln(1+\lambda a^w)$ for $\lambda > 0$ and $w > 0$, of a class of fuzzy complements

$$c_{\lambda,w}(a) = \left(\frac{1-a^w}{1+\lambda a^w}\right)^{1/w}$$

which contains the sugeno class (for $w = 1$) as well as the yager class (for $\lambda = 0$) as a special subclasses.

Theorem 2.8: (2)

Second characterisation theorem of fuzzy complement

Statement:

Let c be a function from $[0,1]$ to $[0,1]$. Then c is a fuzzy complement iff there exists continuous function f from $[0,1]$ to R such that

$f(1) = 0$, f is strictly decreasing and

$c(a) = f^{-1}(f(0) - f(a))$ for all $a \in [0,1]$

Proof.

By the first characterization theorem c is a fuzzy complement iff there exists an increasing generator g such that

$$c(a) = g^{-1}(g(1) - g(a)), \text{ for all } a \in [0,1]$$

Let $f(a) = g(1) - g(a) \rightarrow ①$

Then $f(1) = g(1) - g(1) = 0$

Clearly, f is strictly decreasing since g is strictly increasing.

And $f^{-1}(a) = g^{-1}(g(1) - a)$

$= g^{-1}(f(0) - a) \rightarrow ②$

since $f(0) = g(1) - g(0)$
$= g(1) - 0$
$= g(1)$

$f(f^{-1}(a)) = g(1) - g(f^{-1}(a))$, by ①

$= g(1) - g(g^{-1}(f(0) - a))$

$= g(1) - (f(0) - a)$

$= f(0) - f(0) + a = a$

$f^{-1}(f(a)) = g^{-1}(f(0) - f(a))$ by ②

$= g^{-1}(f(0) - g(1) - g(a))$ by ①

$= g^{-1}(f(0) - f(0) + g(a))$

$= g^{-1}(g(a)) = a$

$\therefore f^{-1}(f(a)) = f(f^{-1}(a))$

$\therefore f$ is continuous

Also, $c(a) = g^{-1}(g(1) - g(a))$

$= f^{-1}[g(a)]$ by ②

$= f^{-1}(g(1) - g(1) + g(a))$

$= f^{-1}(g(1) - (g(1) - g(a)))$

$= f^{-1}[g(1) - f(a)]$ by ①

$= f^{-1}[f(0) - f(a)]$, since $g(1) = f(0)$

$\therefore c(a) = f^{-1}(f(0) - f(a))$

Conversely, If a decreasing generator f is given we can define an increasing generator g as $g(a) = f(0) - f(a)$

Given that $c(a) = f^{-1}(f(0) - f(a))$

$= g^{-1}(g(1) - g(a))$

By the first characterization Theorem, c is fuzzy complement.

Fuzzy Intersections: t - Norms.

Defn: The intersection of two fuzzy sets A and B is defined in general by a function of the form $i : [0,1] \times [0,1] \to [0,1]$. We can write $(A \cap B)(x) = i(A(x), B(x)), \forall x \in X$.

A fuzzy intersection / t-norms i is a binary operation on $[0,1]$ that satisfies atleast the following axioms for all $a, b, d \in [0,1]$.

Axiom (i) : $i(a,1) = a$ (boundary condition)

(ii) : $b \le d$ implies $i(a,b) \le i(a,d)$ (monotonicity)

(iii) : $i(a,b) = i(b,a)$ commutativity

(iv) : $i(a, i(b,d)) = i(i(a,b), d)$ (associativity)

(v) : i is continuous function (continuity)

(vi) : $i(a,a) \le a$ (subidempotency)

(vii) : $a_1 < a_2$ and $b_1 < b_2$ implies $i(a_1, b_1) < i(a_2, b_2)$ (strictly monotonicity)

Result:

1) $i(0,1) = 0$ and $i(1,1) = 1$ from axiom (i)

2) $i(1,0) = 0$ from axiom (iii)

3) $i(0,0) = 0$ from axiom

4) The set of first four axioms are

called automatic section for fuzzy intersection t norms. ... that satisfies

5) A continuous t-norm ... Archimedean t-norm ... subidempotency is called an Archimedean ...

6) A archimedean t-norm satisfies the strict monotonicity, it is called a strictly Archimedean t-norm.

Theorem 2.9:

The standard fuzzy intersection is the only idempotent t-norm.

proof:

Clearly $\min(a,a) = a$, $\forall a \in [0,1]$

Assume that i a t-norm, such that $i(a,a) = a$ $\forall a \in [0,1]$.

To prove that $i(a,b) = \min(a,b)$ $\forall a,b \in [0,1]$.

Now, $a, b \in [0,1]$, $a \leq b$,

$a = i(a,a) \leq i(a,b)$ by axiom (ii)

$\leq i(a,1)$ by "

$= a$ by axiom (i)

$\therefore a \leq i(a,b) \leq a$.

$i(a,b) = a = \min(a,b)$ $\forall a,b \in [0,1]$.

and $a \geq b$,

$b = i(b,b) \leq i(a,b)$ by axiom (ii)

$\leq i(1,b)$ by "

$= b$ by axiom (i)

$b \leq i(a,b) \leq b$.

$b \leq i(a,b) \leq b$, $\forall a,b \in [0,1]$.

$\therefore i(a,b) = b = \min(a,b)$ $\forall a,b \in [0,1]$

$\therefore i(a,b) = \min(a,b)$

Hence the standard fuzzy intersection is the only idempotent t-norm.

Result 1 — t-norm

(1) Some examples of fuzzy intersection / t-norms

a) Standard intersection

$i(a,b) = \min(a,b)$

b) Algebraic product

$i(a,b) = ab$

c) Bounded difference

$i(a,b) = \max(0, a+b-1)$

d) Drastic intersection

$$i(a,b) = \begin{cases} a & \text{when } b = 1 \\ b & \text{when } a = 1 \\ 0 & \text{otherwise} \end{cases}$$

It is denoted by i_{min}

(2) $i_{min}(a,b) \le \max(0, a+b-1) \le ab \le \min(a,b)$

Theorem 2.10: prove that for all $a, b \in [0,1]$, $i_{min}(a,b) \le i(a,b) \le \min(a,b)$, where i_{min} denotes the drastic intersection.

Proof: For all $a, b \in [0,1]$

$i(a,b) \le i(a,1) = a$ by (i)

$i(a,b) \le i(1,b) = b \rightarrow$ (1)

$i(a,b) = i(b,a)$ by (ii)

$i(a,b) \le i(b,1)$

$= b$

$i(a,b) \le b \rightarrow$ (2)

From (1) and (2), $i(a,b) \le \min(a,b) \rightarrow$ (3)

And $i(a,b) = a$ when $b = 1$ but $i(a,1) = a$,

$i(a,b) = b$ when $a = 1$ but $i(1,b) = b$.

$i(a,0) \leq \min(a,0)$, by (3)

$= 0$

$\Rightarrow i(a,0) = 0$,

$i(0,b) \leq \min(0,b)$, by (3)

$= 0$

$\Rightarrow i(0,b) = 0$

$i(a,b) \geq i(a,0)$ by axiom (ii)

$= 0$

$\Rightarrow i(a,b) \geq 0$

$\geq i_{min}(a,b)$

$\Rightarrow i(a,b) \geq i_{min}(a,b)$, where $a,b \in [0,1]$

Hence $i_{min}(a,b) \leq i(a,b) \longrightarrow$ (4)

From (3) and (4), we get

$$i_{min}(a,b) \leq i(a,b) \leq \min(a,b).$$

Defn: A decreasing generator is continuous and strictly decreasing fn from $[0,1]$ to $\mathbb{R}$ such that $f(1) = 0$. The pseudo inverse of a decreasing generator f, denoted by $f^{(-1)}$, is a function from $\mathbb{R}$ to $[0,1]$ given by

$$f^{(-1)}(a) = \begin{cases} 1 & \text{for } a \in (-\infty, 0) \\ f^{-1}(a) & \text{for } a \in [0, f(0)] \\ 0 & \text{for } a \in (f(0), \infty) \end{cases}$$

$(-1) \to$ extension

Where f^{-1} is the ordinary inverse of f.

Example (1): $f_1(a) = 1 - a^p$ for any $a \in [0,1]$ $(p>0)$

is a decreasing generator and

$$f_1^{(-1)}(a) = \begin{cases} 1 & \text{for } a \in (-\infty, 0) \\ (1-a)^{1/p} & \text{for } a \in [0,1] \\ 0 & \text{for } a \in (1, \infty) \end{cases}$$

$y = 1 - a^p$

$a^p = 1 - y$

$a = (1-y)^{1/p}$

$= (1-a)^{1/p}$

is pseudo inverse of f

(ii) $f_0(a) = \ln a$ for any $a \in [0,1]$ with $f_0(0) = \infty$
is a decreasing generator and

$$f_0^{(-1)}(a) = \begin{cases} 1 & \text{for } a \in (-\infty, 0) \\ e^{-a} & \text{for } a \in [0, \infty) \end{cases}$$

is pseudo inverse of f_0.

Note :- A decreasing generator f and its pseudo inverse $f^{(-1)}$ satisfy $f^{-1}(f(a)) = a$ for any $a \in [0,1]$ and

$$f\left(f^{(-1)}(a)\right) = \begin{cases} 0 & \text{for } a \in (-\infty, 0) \\ a & \text{for } a \in [0, f(0)] \\ f(0) & \text{for } a \in (f(0), \infty) \end{cases}$$

Defn :- An increasing generator is a continuous and strictly increasing function g from $[0,1]$ to R such that $g(0) = 0$. The pseudo inverse of an increasing generator denoted by $g^{(-1)}$, is a function from R to $[0,1]$ defined by

$$g^{(-1)}(a) = \begin{cases} 0 & \text{for } a \in (-\infty, 0) \\ g^{-1}(a) & \text{for } a \in [0, g(1)] \\ 1 & \text{for } a \in (g(1), \infty). \end{cases}$$

Where g^{-1} is the ordinary inverse of g.

Example (i) :- $g_1(a) = a^p$, $p > 0$ and $a \in [0,1]$, is a increasing generator and

$$g_1^{(-1)}(a) = \begin{cases} 0 & \text{for } a \in (-\infty, 0) \\ a^{1/p} & \text{for } a \in [0,1] \\ 1 & \text{for } a \in (1, \infty) \end{cases}$$ is

Pseudo inverse of g_1.

② $g_2(a) = -\ln(1-a)$ for any $a \in [0,1]$ with $g_2(1) = \infty$, is a increasing generator and

$$g_2^{(-1)}(a) = \begin{cases} 0 & \text{for } a \in (-\infty, 0) \\ 1-e^{-a} & \text{for } a \in (0,\infty) \end{cases}$$

is pseudo inverse of g_2.

Lemma 2.1: Let f be a decreasing generator. Then a function g defined by $g(a) = f(0) - f(a)$ for any $a \in [0,1]$ is an increasing generator with $g(1) = f(0)$ and its pseudo inverse $g^{(-1)}$ is given by $g^{(-1)}(a) = f^{(-1)}(f(0) - a)$ for any $a \in R$.

Proof: Given that f is a decreasing generator and $g(a) = f(0) - f(a)$. To prove that g is a increasing generator. Clearly g is continuous, strictly decr since f is continuous. For any $a, b \in [0,1]$ $a < b$

$\Rightarrow f(a) > f(b)$ since f is sd

$\Rightarrow f(0) - f(a) < f(0) - f(b)$

$\Rightarrow g(a) < g(b)$

If $a < b$ then $g(a) < g(b)$

$\therefore g$ is strictly increasing.

Also $g(0) = f(0) - f(0) = 0$ and

$g(1) = f(0) - f(1) = f(0) - 0$

$= f(0)$

$g(1) = f(0)$

Hence g is a increasing generator with $g(1) = f(0)$.

As per the definition of increasing generator, the pseudo inverse of g is

$$g^{(-1)}(a) = \begin{cases} 0, & \text{if } a \in (-\infty, 0) \\ g^{-1}(a), & \text{if } a \in [0, g(1)] \\ 1, & \text{if } a \in (g(1), \infty) \end{cases} \longrightarrow (1)$$

$a \in [0, g(1)]$

Let $b = g(a) = f(0) - f(a)$

$\Rightarrow f(a) = f(0) - b$

$\Rightarrow a = f^{-1}(f(0) - b)$

$\Rightarrow g^{-1}(a) = g^{-1}[f^{-1}(f(0) - b)]$

$= f^{-1}[g^{-1}(f(0) - b)]$

$= f^{-1}(f(0) - g^{-1}(b))$

$= f^{-1}(f(0) - a)$

$\Rightarrow g^{-1}(a) = f^{-1}(f(0) - a), \; a \in [0, g(1)]$

i.e. $g^{-1}(a) = f^{-1}(f(0) - a) \quad a \in [0, f(0)] \longrightarrow (2)$

As per the definition of decreasing generator,

$$f^{(-1)}(f(0) - a) = \begin{cases} 1 & \text{if } f(0) - a \in (-\infty, 0) \\ f^{-1}(f(0) - a) & \text{if } f(0) - a \in [0, f(0)] \\ 0 & \text{if } f(0) - a \in (f(0), \infty) \end{cases}$$

$$= \begin{cases} 1 & \text{if } a \in (f(0), \infty) \\ f^{-1}(f(0) - a) & \text{if } a \in [0, f(0)] \\ 0 & \text{if } a \in (-\infty, 0) \end{cases}$$

$$= \begin{cases} 0 & \text{if } a \in (-\infty, 0) \\ g^{-1}(a) & \text{if } a \in [0, f(0) = g(1)] \quad \text{by } (2) \\ 1 & \text{if } a \in (f(0) = g(1), \infty) \end{cases}$$

$= g^{(-1)}(a)$ by ①

$\therefore g^{(-1)}(a) = f^{(-1)}(f(0) - a)$.

Lemma 2.2 :

Let g be an increasing generator. Then the function f defined by $f(a) = g(1) - g(a)$ for any $a \in [0,1]$ is a decreasing generator with $f(0) = g(1)$ and its pseudo inverse $f^{(-1)}$ is given by $f^{(-1)}(a) = g^{(-1)}(g(1) - a)$ for any $a \in R$.

Proof :

Given that g is an increasing generator and $f(a) = g(1) - g(a)$.

To prove that f is decreasing generator.

Clearly f is continuous, since g is continuous.

For any $a, b \in [0,1]$, $a < b$.

$\Rightarrow g(a) < g(b)$

$\Rightarrow g(1) - g(a) > g(1) - g(b)$

$\Rightarrow f(a) > f(b)$

$f(a) > f(b)$ if $a < b$ thus f is strictly decreasing.

Now $f(0) = g(1) - g(0)$
$= g(1) - 0 = g(1)$

$f(0) = g(1)$

and $f(1) = g(1) - g(1) = 0$

$\Rightarrow f$ is decreasing generator with

$f(0) = g(1)$. As per the definition of decreasing generator

$$f^{(-1)}(a) = \begin{cases} 1 & \text{if } a \in (-\infty, 0) \\ f^{-1}(a) & \text{if } a \in [0, f(0)] \\ 0 & \text{if } a \in (f(0), \infty) \end{cases} \rightarrow ①$$

Let $b = f(a) = g(1) - g(a)$, $a \in [0, f(0)]$

$\Rightarrow g(a) = g(1) - b$

$\Rightarrow a = g^{-1}(g(1) - b)$

$\Rightarrow f^{-1}(a) = f^{-1}(g^{-1}(g(1) - b))$

$= g^{-1}(f^{-1}(g(1) - b))$

$= g^{-1}(g(1) - f^{-1}(b))$

For any $a \in [0, f(0)]$ $= g^{-1}(g(1) - a)$ $\Rightarrow$ by (*)

$= [0, g(1)]$

$f^{-1}(a) = g^{-1}(g(1) - a) \rightarrow ②$

As per the definition of increasing generator

$$g^{(-1)}(g(1) - a) = \begin{cases} 0 & \text{if } g(1) - a \in (-\infty, 0) \\ g^{-1}(g(1) - a) & \text{if } g(1) - a \in [0, g(1)] \\ 1 & \text{if } g(1) - a \in (g(1), \infty) \end{cases}$$

$$= \begin{cases} 0 & \text{if } a \in (g(1), \infty) \\ g^{-1}(g(1) - a) & \text{if } a \in [0, g(1)] \\ 1 & \text{if } a \in (-\infty, 0) \end{cases}$$

$$= \begin{cases} 1 & \text{if } a \in (-\infty, 0) \\ f^{-1}(a) & \text{if } a \in [0, f(0)] \text{ by ②} \\ 0 & \text{if } a \in (f(0), \infty) \end{cases}$$

$= f^{(-1)}(a)$ by ①

$\therefore f^{(-1)}(a) = g^{(-1)}(g(1) - a)$

W.A.

Theorem 2.11: Let i_w denote the class of yagor t-norms defined by $i_w(a, b) = 1 - \min(1, [(1-a)^w + (1-b)^w]^{1/w})$, $w > 0$. Thus $i_{min}(a, b) \leq i_w(a, b) \leq \min(a, b)$ $\forall\, a, b \in [0, 1]$.

Proof: Given that

$$i_w(a,b) = 1-\min\left(1, \left[(1-a)^w + (1-b)^w\right]^{1/w}\right)$$

Put $a=1$ we get

$$i_w(1,b) = 1-\min(1, 1-b)$$
$$= 1-(1-b)$$
$$= 1-1+b = b$$

Put $b=1$ we get

$$i_w(a,1) = 1-\min(1, 1-a)$$
$$= 1-1+a$$
$$= a$$

(i) ∴ $i_{\min}(a,1) = i_w(a,1)$ and

$i_{\min}(1,b) = i_w(1,b)$

Lt $\left[(1-a)^w + (1-b)^w\right]^{1/w} = (1+1)^{1/0} = 2^{\infty} = \infty$

$w \to 0$

Lt $i_w(a,b) = 1-\min(1, \infty)$

$w \to 0$

$= 1-1$

$= 0$

for $a,b \in [0,1]$

$i_{\min}(a,b) \le i_w(a,b) \Rightarrow$ ①

Lt $(a^w+b^w)^{1/w} = \max(a,b)$

$w \to \infty$

(ii) Lt $\min\left(1, \left((1-a)^w + (1-b)^w\right)^{1/w}\right) = \max(1-a, 1-b)$

$w \to \infty$

$i_\infty(a,b) = 1-\max(1-a, 1-b)$ by demorgan

$(A^c \cup B^c)^c$

$= A^{cc} \cap B^{cc}$

$= A \cap B$

$= \min$

$= \min(a,b)$

$i_w(a,b) \le \min(a,b) \Rightarrow$ ②

from ① and ②

$i_{\min}(a,b) \le i_w(a,b) \le \min(a,b)$

Fuzzy unions : t - conorms

Defn : The union of two fuzzy sets A and B is specified in general by a function of the form $u : [0, 1] \times [0, 1] \rightarrow [0, 1]$ we can write

$(A \cup B)(x) = u(A(x), B(x))$ for all $x \in X$.

A fuzzy union / t - conorm u is a binary operation on the unit interval that satisfies atleast the following axiom for all $a, b, d \in [0,1]$.

Axiom (i) : $u(a, 0) = a$ (boundary condition)

Axiom (ii) : $b \le d$ implies $u(a,b) \le u(a,d)$ (monotonicity)

Axiom (iii) : $u(a,b) = u(b,a)$ commutativity.

Axiom (iv) : $u(a, u(b,d)) = u(u(a,b), d)$, associativity

Axiom (v) : u is a continuous function (continuity)

Axiom (vi) : $u(a,a) > a$ (super idem potency)

Axiom (vii) : $a_1 < a_2$ and $b_1 < b_2$ implies $u(a_1, b_1) < u(a_2, b_2)$ (Strictly monotonicity)

Result : 1) $u(0,1) = 1$ and $u(1,1) = 1$ from axiom (i)

2) $u(1,0) = 1$ from axiom (iii)

3) $u(0,0) = 0$ from axiom (i)

4) The set of first four axioms are called axiomatic skelton for fuzzy union, t-conorms.

5) A continuous t-conorm that satisfies super idem potency is called an Archimedean t-conorm.

6) A archimedian t-conorm satisfies the strictly monotonicity, it is called a strictly Archimedean t-conorm.

Theorem 2.14 :

The standard fuzzy union is the only idempotent t-conorm.

proof : W.k.T max$(a,a) = a$ $\forall a \in [0,1]$

Assume that $\exists$ a t-conorm u such that $u(a,a) = a$ $\forall a \in [0,1]$

To prove that $u(a,b) = \min(a,b)$ $\forall (a,b) \in [0,1]$

Now

$a, b, d \in [0,1]$; $(a \geq b)$

$a = u(a, 0) \geq u(a, b)$ (by axiom (ii))

$\geq u(a, 0)$

$= a$ by axiom (i)

$\therefore a \geq u(a, b) \geq a$ $\forall a, b \in [0,1]$

$\therefore u(a, b) = a = \min(a, b)$

and $(a \leq b)$

$b = u(b, b) \geq u(a, b)$ by axiom (ii)

$\geq u(0, b)$

$= b$ (by axiom (i))

$b \geq i(a, b) \geq b$

$u(a, b) = b = \min(a, b)$ $\forall a, b \in [0,1]$

$\therefore u(a, b) = \min(a, b)$ $\forall a, b \in [0,1]$

Hence the standard fuzzy union is the only idempotent t-conorm.

Result

ii) Some examples of fuzzy union / t-conorm

a) Standard union

$u(a, b) = \max(a, b)$

b) Algebraic sum

$u(a, b) = a + b - ab$

c) bounded difference

$u(a, b) = \min(1, a + b)$

d) Drastic union

$$u(a, b) = \begin{cases} a & \text{when } b = 0 \\ b & \text{when } a = 0 \\ 1 & \text{otherwise} \end{cases}$$

It is denoted ...

2) $\max(a,b) \le a+b-ab \le \min(1, a+b) = u_{max}(a,b)$

Theorem 2.15:

For all $a, b \in [0,1]$, prove that $\max(a,b) \le u(a,b) \le u_{max}(a,b)$, where u_{max} denotes the drastic union.

Proof: For all $a, b \in [0,1]$

$u(a,b) \ge u(a,0)$ by axiom C(ii)

$= a$ by (i)

$\therefore u(a,b) \ge a \rightarrow (1)$

and $u(a,b) = u(b,a)$ by (iii)

$\ge u(b,0)$ by (ii)

$= b$

$\therefore u(a,b) \ge b \rightarrow (2)$

From (1) and (2)

$\max(a,b) \le u(a,b)$

And $u(a,b) = a$ when $b=0$ but $u(a,0) = a$.

$u(a,b) = b$ when $a=0$ but $u(0,b) = b$.

$u(a,1) \ge \max(a,1)$ by (3)

$\cancel{u(a,b) =}$

$u(0,b) \ge \max(0,b)$ by (3)

$= 0$

$u(1,b) = 1$

$u(a,b) \ge u(a,1)$ by axiom C(ii)

$= 1$

$\Rightarrow u(a,b) \le 1.$

$\le u_{max}(a,b)$

$u(a,b) \le u_{max}(a,b)$ where $a, b \in [0,1]$.

Hence $u_{max}(a,b) \ge u(a,b) \rightarrow (4)$

From (3) and (4) we get

$\max(a,b) \le u(a,b) \le u_{max}(a,b).$

Theorem 3.16

Let u_w denote the class of Yager u- defined by

$$u_w(a,b) = \min\left(1, (a^w + b^w)^{1/w}\right),$$

$w > 0$. Then, $\max(a,b) \le u_w(a,b) \le u_{max}(a,b)$ for all $a, b \in [0,1]$.

Proof: Let $a, b \in [0,1]$

First we prove that

$$\lim_{w \to \infty} u_w(a,b) = \max(a,b)$$

__Case (i)__ If $a = 0$, then

$$\lim_{w \to \infty} u_w(0,b) = \lim_{w \to \infty} \min(1,b) = b$$

and $\max(a,b) = \max(0,b) = b$ if $a = 0$.

$$\therefore \lim_{w \to \infty} u_w(a,b) = \max(a,b)$$

__Case (ii)__ If $b = 0$, then

$$\lim_{w \to \infty} u_w(a,0) = \lim_{w \to \infty} \min(1,a) = a$$

and $\max(a,b) = \max(a,0) = a$ if $b = 0$.

$$\therefore \lim_{w \to \infty} u_w(a,b) = \max(a,b)$$

__Case (iii)__ If $a = b$, then

$$\lim_{w \to \infty} u_w(a,b) = \lim_{w \to \infty} \min\left(1, 2^{1/w} a\right)$$

$$= \min(1,a) \quad \text{since } \lim_{w \to \infty} 2^{1/w} = 1$$

$$= a = \max(a,a)$$

and $\max(a,b) = \max(a,b)$ if $a = b$.

Case (iv): If $a \neq b$ is then clearly

$$\min\left(1, (a^w + b^w)^{1/w}\right) = (a^w + b^w)^{1/w}$$

It is enough to prove that

$$\lim_{w \to \infty} \left(a^w + b^w\right)^{1/w} = \max(a, b)$$

Let us assume, without loss of generality, and if $a < b$, then

$$Q = (a^w + b^w)^{1/w}$$

$$\Rightarrow \ln Q = \frac{1}{w} \ln (a^w + b^w)$$

$$\Rightarrow \lim_{w \to \infty} \ln Q = \lim_{w \to \infty} \frac{\ln (a^w + b^w)}{w} = \frac{\infty}{\infty}$$

Using L-Hospital rule,

$$\therefore \lim_{w \to \infty} \ln Q = \lim_{w \to \infty} \frac{\frac{1}{a^w + b^w}\left(a^w \ln a + b^w \ln b\right)}{1}$$

$$= \lim_{w \to \infty} \frac{a^w \ln a + b^w \ln b}{a^w + b^w}$$

$$= \lim_{w \to \infty} \frac{\left(\frac{a}{b}\right)^w \ln a + \ln b}{\left(a/b\right)^w + 1}$$

$= \ln b$, since $\lim_{w \to \infty} (a/b)^w = 0$

$\lim_{w \to \infty} Q = b$ and

$$\max(a, b) = b$$

$$\lim_{w \to \infty} \left(a^w + b^w\right)^{1/w} = \max(a, b)$$

Hence $\lim_{w \to \infty} \min\left(1, (a^w + b^w)^{1/w}\right) = \max(a, b)$

When $\min\left(1, (a^w + b^w)^{1/w}\right) = 1$ in the case $(a^w + b^w)^{1/w} \geq 1 \Rightarrow (a^w + b^w) \geq 1$,

for all $w \in (0, \infty)$

(when $w \to \infty$) the last inequality holds if $a = 1$ or $b = 1$

$\therefore \max(a,b) = \max(a,b)$

$\therefore \lim_{\omega \to \infty} \min(1, (a^{\omega}+b^{\omega})^{1/\omega}) = \max(a,b) \longrightarrow ①$

$\therefore \max(a,b) \le u_{\omega}(a,b)$

$u_{\omega}(a,b) \le u_{max}(a,b)$.

Case (a) : If $a=0$, then

$u_{\omega}(0,b) = \min(1,b) = b$

and $u_{max}(0,b) = b$

$\therefore u_{\omega}(a,b) = u_{max}(a,b)$ if $a=0$

Case (b). If $b=0$, then

$u_{\omega}(a,0) = \min(1,a) = a$

and $u_{max}(a,0) = a$

$\therefore u_{\omega}(a,b) = u_{max}(a,b)$ if $b=0$

Case (c) otherwise

$\therefore \lim_{\omega \to 0} \min(1, (a^{\omega}+b^{\omega})^{1/\omega}) =$

$= \min(1,\infty)$

$= 1$ $a,b \in (0,1]$

and $u_{max}(a,b) = 1$, for all $a,b \in [0$

$\therefore u_{\omega}(a,b) \le u_{max}(a,b)$

From ① and ②, we get $u_{max}(a,b)$

$u_{max}(a,b) \le u_{\omega}(a,b)$

UNIT - III

Combinations of operations :

Defn :

A t-norm i and a t-conorm u are dual with respect to a fuzzy complement c iff

$c(i(a,b)) = u(c(a), c(b))$ and $c(u(a,b)) = i(c(a), c(b))$

These equations describes the De-morgan's laws for the fuzzy sets, let the triple $\langle i, u, c\rangle$ denote that i and u are dual with respect to c and let any such triple $\langle i, u, c\rangle$ be called a dual triple.

Result : Some dual t-norms and t-conorms with respect to the statement complement c,

i) $\langle \min(a,b), \max(a,b), c_s \rangle$

ii) $\langle ab, a+b-ab, c_s \rangle$

iii) $\langle \max(0, a+b-1), \min(1, a+b)$

iv) $\langle i_{\min}(a,b), u_{\max}(a,b), c_s \rangle$

Theorem 3.17 :

The triple $\langle min, max, c \rangle$ and $\langle i_{min}, u_{max}, c \rangle$ are dual with respect to any fuzzy complement c.

proof :

Assume, without loss of generality, that $a \leq b$

$\Rightarrow c(a) \geq c(b)$.

i) $max(c(a), c(b)) = c(a)$ and

$c(min(a, b)) = c(a)$

$\therefore max(c(a), c(b)) = c(min(a, b))$, for all $a, b \in [0,1]$

ii) $min(c(a), c(b)) = c(b)$ and

$c(max(a, b)) = c(b)$

$min(c(a), c(b)) = c(max(a, b)) \ \forall a, b \in [0,1]$

$\therefore \langle min, max, c \rangle$ is a dual triple.

iii) Case (i) if $a = 0$, $b \in [0,1]$, then

$c(a) = c(0) = 1 \geq c(b)$

$\therefore u_{max}(c(a), c(b)) = u_{max}(1, c(b)) = 1$

$c(i_{min}(a, b)) = c(0) = 1$

$\therefore u_{max}(c(a), c(b)) = c(i_{min}(a, b))$ if $a = 0$.

& $u_{min}(c(a), c(b)) = u_{min}(1, c(b)) = c(b)$.

$c(i_{min}(a, b)) = c(b)$

$\therefore u_{min}(c(a), c(b)) = c(i_{min}(a, b))$ if $a = 1$.

Case ii) if $b = 0$, $a \in [0,1]$,

then $c(b) = c(0) = 1 \geq c(a)$

By case i), $u_{max}(c(a), c(b)) = c(i_{min}(a, b))$ if $b = 0$

$u_{max}(c(a), c(b)) = c(i_{max}(a, b))$ if $a = 1$

$u_{min}(c(a), c(b)) = c(i_{max}(a, b))$ (by case i)

$c[0,1]$

Case (iii) If $a = 1$; $b \in [0,1)$, then

$c(a) = c(1) = 0 \leq c(b)$,

$\therefore u_{max}(c(a), c(b)) = u_{max}(0, c(b))$
$= c(b)$

and $c(i_{min}(a,b)) = c(i_{min}(1,b))$
$= c(b)$.

$\therefore u_{max}(c(a), c(b)) = c(i_{min}(a,b))$ if $a = 1$.

And $i_{min}(c(a), c(b)) = i_{min}(0, c(b))$
$= 0$,

$c(u_{max}(a,b)) = c(u_{max}(1,b))$
$= c(1) = 0$

$\therefore i_{min}(c(a), c(b)) = c(u_{max}(a,b))$ if $a=1$, then

Case (iv) : If $b = 1$; $a \in [0,1)$,

$c(b) = c(1) = 0 \leq c(a)$,

by case (iii),

$u_{max}(c(a), c(b)) = c(i_{min}(a,b))$

and $i_{min}(c(a), c(b)) = c(u_{max}(a,b))$

Case (v) : Otherwise, if $a \leq b$, then

$c(a) \geq c(b)$,

$u_{max}(c(a), c(b))$

$c(i_{min}(a,b)) = c(0)$

$\therefore u_{max}(c(a), c(b)) = c(i_{min}(a,b))$

$i_{min}(c(a), c(b)) = 0$

$c(u_{max}(a,b)) = 0$

$i_{min}(c(a), c(b)) = c(u_{max}(a,b))$

$\therefore \langle i_{min}, u_{max}, c \rangle$ is a dual triple.

Theorem 2.18 :-

Given a t-norm i and an involutive fuzzy complement c, the binary operation u on $[0,1]$ defined by $u(a,b) = c(i(c(a), c(b)))$ for all $a, b \in [0,1]$ is a t-conorm such that $\langle i, u, c\rangle$ is a dual triple.

Proof :-

First we prove that u is a t-conorm.

For any $a, b \in [0,1]$, then

$$u(a,0) = c(i(c(a), c(0)))$$
$$= c(i(c(a), 1))$$
$$= c(c(a)) = a$$

Axiom (i) is satisfied.

For any $a, b, d \in [0,1]$

if $b \leq d$, then $c(b) \geq c(d)$

also $i(c(a), c(d)) \leq i(c(a), c(b))$

$$c(i(c(a), c(d))) \geq c(i(c(a), c(b))) \rightarrow (1).$$

$$u(a,b) = c(i(c(a), c(b)))$$
$$\leq c(i(c(a), c(d))) \text{ by } (1).$$
$$= u(a,d)$$

$u(a,b) \leq u(a,d)$ $\forall a, b, d \in [0,1]$

Axiom (ii) is satisfied.

For any $a, b \in [0,1]$

$$u(a,b) = c(i(c(a), c(b)))$$
$$= c(i(c(b), c(a)))$$
$$= u(b,a) \quad \forall a, b \in [0,1].$$

Axiom (iii) is satisfied.

For any $a, b, d \in [0,1]$

$$u(a, u(b,d)) = c[i(c(a), c(u(b,d)))]$$
$$= c[i(c(a), c[c(i(c(b), c(d)))])]$$

$= c[i(c(a), i(c(b), c(d)))]$

$= c[i(c(a), c(b))], c(d))$

$= c[i(c(i(c(a), c(b)))), c(d)]$

$= c[i(c(c(i(c(a), c(b))))), c(d)]$

$= c[i(c(u(a,b))), c(d))$

$= u(u(a,b), d)$

$\therefore u(a, u(b,d)) = u(u(a,b), d) \quad \forall a,b,d \in [0,1]$

$\therefore$ Axiom (iv) is satisfied.

Next we prove that $\langle i, u, c \rangle$ is a dual triple.

$c(u(a,b)) = c[c(i(c(a), c(b)))]$

$= i(c(a), c(b))$

$\therefore c(u(a,b)) = i(c(a), c(b))$ — (1)

$u(c(a), c(b)) = c[i(c(c(a)), c(c(b)))] = c(i(a,b))$

$\therefore c[i(a,b)] = u(c(a), c(b))$

$\therefore \langle i, u, c \rangle$ is a dual triple.

Theorem

Given a t-conorm u and an involutive fuzzy complement c, the binary operation i on $[0,1]$ defined by $i(a,b) = c(u(c(a), c(b)))$ for all $a, b \in [0,1]$ is a t-norm such that $\langle i, u, c \rangle$ is a dual triple.

Proof: First we prove that i is a t-norm.

for any $a \in [0,1]$, then

$i(a,1) = c(u(c(a), c(1)))$

$= c(u(c(a), 0))$

$\therefore c(c(a)) = a$

Axiom c(ii) is satisfied.

For any $a, b, d \in [0,1]$,

if $b \le d$, then $c(b) \ge c(d)$

also $u(c(a), c(d)) \ge u(c(a), c(b))$

$c(u(c(a), c(d)) \le c(u(c(a), c(b))) \rightarrow (1)$

$i(a,b) = c(u(c(a), c(b)))$

$\le c(u(c(a), c(d))$ by (1)

$\le i(a,d)$.

$i(a, b) \le i(a,d) \; \forall \; a, b, d \in [0,1]$

$\therefore$ Axiom c(ii) is satisfied.

For any $a, b \in [0,1]$

$i(a,b) = c(u(c(a), c(b)))$

$= c(u(c(b), c(a)))$

$= i(b,a), \; \forall \; a, b \in [0,1]$.

Axiom c(iii) is satisfied.

For any $a, b, d \in [0,1]$.

$i(a, i(b,d)) = c[u(c(a), c[i(b,d)])]$.

$= c[u(c(a), c[c[u(c(b), c(d)$

$= c[u(c(a), u(c(b), c(d)))]$

$= c[u(u(c(a), c(b)), c(d)))$.

$= c[u[c(c(u(c(a), c(b)))), c(d)]]$

$= c[u[c(i(a,b))], c(d))]$

$= i(i(a,b), d)$

$i(a, i(b,d)) = i(i(a,b), d)$

$\forall a, b, d \in [0,1]$

axiom c(iv) is satisfied.

Next we prove that $\langle u, i, c \rangle$ is a dual triple.

$$c(i(a,b)) = c(c(u(c(a), c(b))))$$

$$= u(c(a), c(b))$$

$$\therefore c(i(a,b)) = u(c(a), c(b))$$

$$i(c(a), c(b)) = c(u(c(c(a)), c(c(b))))$$

$$= c(u(a,b))$$

$$\therefore c(u(a,b)) = i(c(a), c(b))$$

$\therefore \langle u, i, c \rangle$ is a dual triple.

∴ A ≠ B is

Fuzzy Relations :-

Defn : A relation among crisp sets $x_1, x_2, \ldots$ x_n is a subset of the cartesian product

$X_1 \cdot X_2 \cdots$ is denoted $\cdots X_n$

i.e. N_n

and $R(x_1, x_2, \ldots, x_n) \subset X_1 \times X_2 \times \cdots \times X_n$

The characteristic function of the relation is defined as

$$R(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & (x_1, x_2, \ldots, x_n) \in R \\ 0 & \text{otherwise.} \end{cases}$$

Defn: A relation between two sets is called binary.

Defn: A relation between three sets is called ternary.

Defn: A relation between four sets is called quaternary.

Defn: A relation between five sets is called quinary.

Defn: A relation between n-sets is called n-ary or n-dimension.

Example: Let R be a relation among the three sets X = {English, French} and

Y = {dollar, pound, franc, mark},

Z = {US, France, Canada, Britain, Germany},

which associates a country with a currency and language as follows:

$R(X, Y, Z) = \{$ ⟨English, dollar, US⟩, ⟨French, France, franc⟩,

< English, dollar, Canada >,
< French, dollar, Canada >,
< English, pound, Britains }.

Soln:-

	US	fra	Can	Brit	Ger
dollar	1	0	1	0	0
pound	0	0	0	1	0
Franc	0	0	0	0	0
mark	0	0	0	0	0

English

	US	fra	Can	Brit	Ger
dollar	0	0	0	0	0
pound	0	0	0	0	0
franc	0	1	0	0	0
mark	0	0	0	0	0

Def: Let X and Y be two crisp set. The fuzzy relation R is defined as $R : X \times Y \to [0,1]$

Example: Let R be a fuzzy relation between two sets X = {New york city, paris} and Y = {Beijing, New york city, London} that represents the relational concept "very far". the fuzzy relation is defined as:

$R(x, y) = \{1/$ NY city, Beijing $+ 0/$ NYC, NYC $+ 0.6/$ NYC, London $+ 0.9/$ paris, Beijing $+ 0.7/$ paris, NYC $+ 0.3/$ paris, London}

Find the two-dimensional membership array of

Soln: The two dimensional membership array for the fuzzy relation R is

$$\begin{array}{c|c} & \text{NYC} \quad \text{paris} \\ \text{Beijing} & \\ \text{NYC} & \\ \text{London} & \end{array} \begin{bmatrix} 1 & 0.9 \\ 0 & 0.7 \\ 0.6 & 0.3 \end{bmatrix}$$

Def: Let $X = \{X_i / i \in N_n\}$ be a collection of sets. Then the cartesian product $\times_{i \in N_n} X_i$, whose elements are $x = (x_i / i \in N_n) \in \times_{i \in N_n} X_i$. Thus the $y = (y_j / j \in J) \in \times_{j \in J} X_j$, $J \subseteq N_n$ is called a subsequence of x, where $x_j = y_j \; \forall j \in J$.

projection

Def: Consider the Relation $R(X_1, X_2, \ldots X_n)$. Then the projection of R is $[R \downarrow Y]$ that disregards all variables in X except those in the set $Y = \{X_j / j \in J \subset N_n\}$. Thus $[R \downarrow Y]$ is a fuzzy set whose membership function is defined on the cartesian product of sets in Y by the equation

$M_{[R \downarrow Y]}(y) = \max M_R(x)$, where $x \succ y$, x is the extension. M_R is the membership function of the given n-ary relation R.

$[R \downarrow Y]$ a subset of Y

Example: Consider the sets $X_1 = \{x, y\}$, $X_2 = \{a, b\}$ and $X_3 = \{*, \$\}$ and the ternary fuzzy relation

$$R(x_1, x_2, x_3) = \frac{0.9}{x,a,*} + \frac{0.4}{x,b,*} + \frac{1}{y,a,*} + \frac{0.7}{y,a,\$} + \frac{0.8}{y,b,\$}$$

defined on $X_1 \times X_2 \times X_3$. Find $R_{ij} = [R \downarrow \{X_i, X_j\}]$ and $R_i = [R \downarrow \{X_i\}]$ for all $i, j \in N_3$.

Soln :-

$$R_{1,2} = \frac{0.9}{x,a} + \frac{0.4}{x,b} + \frac{1}{y,a} + \frac{0.8}{y,b}$$

$$R_{1,3} = \frac{0.9}{x,*} + \frac{0}{x,\$} + \frac{1}{y,*} + \frac{0.8}{y,\$}$$

$$R_{2,3} = \frac{1}{a,*} + \frac{0.7}{a,\$} + \frac{0.4}{b,*} + \frac{0.8}{b,\$}$$

$$R_1 = \frac{0.9}{x} + \frac{1}{y}$$

$$R_2 = \frac{1}{a} + \frac{0.8}{b}$$

$$R_3 = \frac{1}{*} + \frac{0.8}{\$}$$

Defn :- Let R be a relation defined on the Cartesian product of sets in the family Y, and let $[R \uparrow X - Y]$ denote the cylindric extension of R into sets X_i $(i \in N_n)$ that are in X but are not in Y. Then $[R \uparrow X - Y]$ is a fuzzy set whose membership function is defined on X:

$$\mu_{[R \uparrow X - Y]}(x) = \mu_R(y)$$

for each x such that $x \succ y$.

Example : Consider the sets $X_1 = \{x, y\}$, $X_2 = \{a, b\}$ and $X_3 = \{*, \$\}$. Find the membership function of cylindric extension of

$$R_{1,2} = \frac{.9}{x,a} + \frac{0.4}{x,b} + \frac{1}{y,a} + \frac{.8}{y,b}$$

$$R_{1,3} = \frac{\cdot 9}{x,*} + \frac{1}{y,*} + \frac{\cdot 8}{y,\$}$$

$$R_{2,3} = \frac{1}{a,*} + \frac{\cdot 4}{b,*} + \frac{\cdot 7}{a,\$} + \frac{0\cdot 8}{b,\$}$$

$$R_1 = \frac{\cdot 9}{x} + \frac{1}{y}$$

$$R_2 = \frac{1}{a} + \frac{\cdot 8}{b}$$

$$R_3 = \frac{1}{*}, \frac{\cdot 8}{\$}$$

	$R_{1,2}$	$R_{1,3}$	$R_{2,3}$	R_1	R_2	R_3
(x, a, *)	.9	.9	1	.9	1	1
(x, a, \$)	.9	0	.7	.9	1	.8
(x, b, *)	.4	.9	.4	.9	.8	1
(x, b, \$)	.4	0	.8	.9	.8	.8
(y, a, *)	1	1	1	1	1	1
(y, a, \$)	1	.8	.7	1	1	.8
(y, b, *)	.8	[illegible]	.4	1	.8	1
(y, b, \$)	.8	.8	.8	1	.8	.8

Defn

The cyclindric closure of the projection $\{R_i / i \in I\}$ is defined as

$$\mu_{cyl} \{R_i\}(x) = \min_{i \in I} \mu [R_i \uparrow x - y_i](x)$$

Example:

Find the cylindric closure of the projection $R_{1,2}$, $R_{1,3}$, $R_{2,3}$ (using the previous example).

soln: cylindric closure

$$cyl\{R_{1,2}, R_{1,3}, R_{2,3}\} = \frac{\cdot 9}{x, a, *} + \frac{[illegible]}{x, b, *} + \frac{1}{y, a, *} + \frac{0.7}{y, a, \$} + \frac{0.4}{y, b, *} + \frac{\cdot 8}{y, b, \$}$$

Binary Relation:

Let X and Y be any two sets. The binary relation R is a subset of $X \times Y$.

i.e. $R = \{(x,y) / x \in X \text{ and } y \in Y\}$.

Domain of R:

The Domain of R is defined by

$Dom(R) = \{x / (x,y) \in R\}$

Range of R:

The Range of R is defined by

$Range(R) = \{y / (x,y) \in R\}$

Fuzzy binary relation:

Let X and Y be any two sets. The fuzzy binary relation (or) fuzzy relation $R(x,y)$ is defined as

$R(x,y) = \{\langle (x,y), \mu_R(x,y) \rangle / x \in X \text{ and } y \in Y\}$

$X \times Y \to [0,1]$.

Defn: Let $R(x,y)$ be a fuzzy relation. The domain of R is a fuzzy set dom $R(x,y)$, where the membership value of dom $R(x,y)$ is defined as

$\mu_{dom R}(x) = \max_{y \in Y} \mu_R(x,y)$ for each $x \in X$.

Defn: If $R(x,y)$ is a fuzzy relation, its range is the fuzzy set ran $R(x,y)$ where the membership value of ran $R(x,y)$ is defined as

$\mu_{ran R}(y) = \max_{x \in X} \mu_R((x,y))$ for each $y \in Y$.

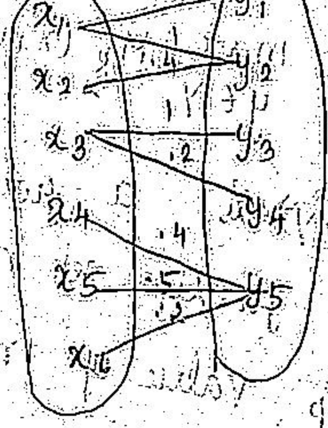
Defn: Let $R(x,y)$ be a fuzzy relation. The height $h(R)$ of the fuzzy relation R is a number defined by
$$h(R) = \max_{y \in Y} \max_{x \in X} \mu_R(x,y).$$

Def: If $h(R) = 1$, then the fuzzy relation is called a normal otherwise it is called subnormal.

Def: Let R be a binary fuzzy relation on sets X and Y. The domain of R is equal to the support of the set X, then the relation is called completely specified, otherwise it is called incompletely specified.

Def: Let R be a binary fuzzy relation. If the range of R is equal to the support of the set Y, then R is called a relation from X onto Y, otherwise it is called a relation from X into Y.

Example: consider the following fuzzy relation R,



Find (i) domain of R (ii) Range of R (iii) verify the relation R is completely specified (a.. point) and onto (preimage). (normal.. one 1)
(iv) verify the relation R is normal and function (mapping). function - no 1 image
Find (i)-(iv)

Soln :- The domain of R is

domain $R = \{(x_1, 1), (x_2, .4), (x_3, 1), (x_4, .4), (x_5, .5), (x_6, .2)\}$

ii) The range of R is

$ran(R) = \{(y_1, .9), (y_2, 1), (y_3, 1), (y_4, .2), (y_5, .5)\}$

iii) Clearly R is completely specified and onto

iv) Clearly R is normal since $h(R) = 1$. but R is not a function, since the range of x_1 is y_1 and y_2.

Note: Every fuzzy relation $R(x, y)$ can be represented by its resolution form

$$R = \bigcup_{\alpha} \alpha R_{\alpha}, \quad \alpha \in \Lambda_R \text{ (level set of } R)$$

where αR_α is a fuzzy relation defined by $\mu_{\alpha R_\alpha}(x, y) = \alpha \cdot \mu_{R_\alpha}(x, y)$ for every $(x, y) \in X \times Y$

Example: Let a binary fuzzy relation R be defined by the following membership

$$\mu_R = \begin{bmatrix} .7 & .4 & 0 \\ .9 & 1 & .4 \\ 0 & .7 & 1 \\ .7 & .9 & 0 \end{bmatrix}$$

Find the resolution from of R.

Soln: The level set of R is

$$\Lambda_R = \{0, .4, .7, .9, 1\}$$

If $\alpha = 0$, then $R_\alpha = \phi$

The resolution form of R is

$$\mu_R = .4 \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cup .7 \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cup .9 \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cup 1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} .4 & .4 & 0 \\ .4 & .7 & .4 \\ 0 & .4 & .4 \\ .4 & .4 & 0 \end{bmatrix} \cup \begin{bmatrix} .4 & 0 & 0 \\ .7 & .7 & 0 \\ 0 & .4 & .4 \\ .7 & .7 & 0 \end{bmatrix} \cup \begin{bmatrix} 0 & 0 & 0 \\ .9 & .7 & 0 \\ 0 & 0 & .7 \\ 0 & .7 & 0 \end{bmatrix} \cup \Big[\text{[illegible]}$$

max

$$= \begin{bmatrix} .7 & .4 & 0 \\ .9 & 1 & .4 \\ 0 & .7 & 1 \\ .7 & .7 & 0 \end{bmatrix}$$

$$= \mu_R$$

Def:

The inverse of a crisp relation $R(x,y)$ is defined as

$R^{-1}(x,y) = \{ (y,x) / (x,y) \in R \} \ \forall \ x \in X$ and $y \in Y$.

Def:

Let $R(x,y)$ be a fuzzy relation then the inverse of $R(x,y)$ is defined by

$R^{-1}(x,y) = \{ ((x,y), \mu_{R^{-1}}(x,y)) / \mu_{R^{-1}}(x,y) = \mu_R(y,x) \}$

Note:

$\mu_{R^{-1}}(x,y)$ - transpose of the $\mu_R(x,y)$.

Example:

$R(x,y)$ be a fuzzy relation on $X = \{x,y,z\}$ and $Y = \{a,b\}$ such that

$$\mu_R = \begin{array}{c|c} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} x \\ y \\ z \end{array} & \begin{bmatrix} .3 & .2 \\ 0 & 1 \\ .6 & .4 \end{bmatrix} \end{array}$$

Find the inverse of $R(x,y)$.

Soln: The inverse of $R(x,y)$ is $R^{-1}(x,y)$ whose membership matrix is

$$\mu_{R^{-1}} = \begin{array}{c|c} & \begin{array}{ccc} x & y & z \end{array} \\ \begin{array}{c} a \\ b \end{array} & \begin{bmatrix} .3 & 0 & .6 \\ .2 & 1 & .4 \end{bmatrix} \end{array}$$

Defn: The composition of any two fuzzy relation $P(x,y)$ and $Q(y,z)$ is defined as $P(x,y) \circ Q(x,z)$

is defined as

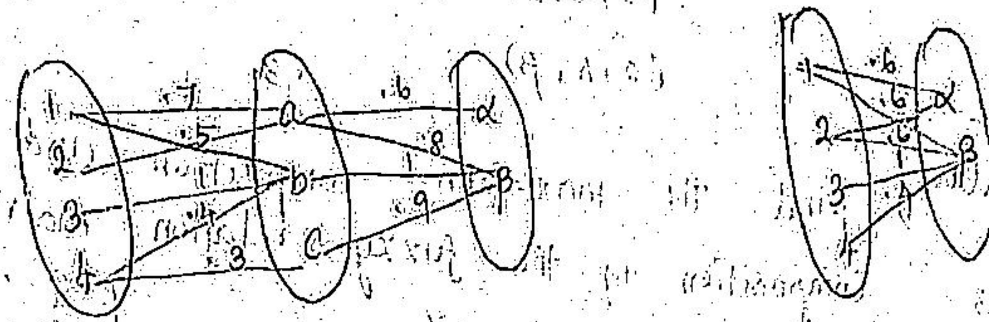
$\mu_{P\circ Q}(x,z) = \max_{y\in Y} \min\{\mu_P(x,y), \mu_Q(y,z)\}$

$\forall\, x\in X,\ z\in Z$. It is denoted as max-min composition.

Example.

Consider the two binary fuzzy relation $P(X,Y)$ and $Q(Y,Z)$ is sagittal diagram.

The composition of P and Q is



Def: The max product composition of any two fuzzy relation $P(X,Y)$ and $Q(Y,Z)$ is defined as $P(X,Y) \odot Q(Y,Z)$ where membership value is defined as $\mu_{P\odot Q} = \max_{y\in Y}\{\mu_P(x,y)\cdot\mu_Q(y,z)\}$

$\forall\, x\in X$ and $z\in Z$

Note: $\odot \to *$

Def: The Joint of any two fuzzy relation $P(X,Y)$ and $Q(Y,Z)$ is defined as $P(X,Y)*Q(Y,Z)$ whose membership value is defined as

$\mu_{P*Q}(x,y,z) = \min\{\mu_P(x,y), \mu_Q(y,z)\}$, $\forall\, x\in X$, $y\in Y$ and $z\in Z$.

Ex: By the previous example

Sol:

Points (x, y, z)	$\mu_S(x, y, z)$
$(1, a, \alpha)$	.6
$(1, a, \beta)$	.7
$(1, b, \alpha)$	.5
$(1, b, \beta)$	.6
$(2, a, \alpha)$	.8
$(2, a, \beta)$	1
$(2, b, \alpha)$	.4
$(2, b, \beta)$	.3

1) Find the max-min composition and max-prod composition of the fuzzy relation matrices

$$\mu_P = \begin{bmatrix} .3 & .5 & .8 \\ 0 & .7 & 1 \\ .4 & .6 & .5 \end{bmatrix} \text{ and } \mu_Q = \begin{bmatrix} .9 & .5 & .7 & .7 \\ .3 & .2 & 0 & .9 \\ 1 & 0 & .5 & .5 \end{bmatrix}$$

The max-min composition of the fuzzy relation P and Q is

$$\mu_{P \odot Q} = \begin{bmatrix} .3 & .5 & .8 \\ 0 & .7 & 1 \\ .4 & .6 & .5 \end{bmatrix} \odot \begin{bmatrix} .9 & .5 & .7 & .7 \\ .3 & .2 & 0 & .9 \\ 1 & 0 & .5 & .5 \end{bmatrix}$$

$$= \begin{bmatrix} \max\{.3, .3, .8\} & \max\{.3, .2, 0\} & \max\{.3, 0, .5\} & \max\{.3, .5, [illegible]\} \\ \max\{0, .3, 1\} & \max\{0, .2, 0\} & \max\{0, 0, .5\} & \max\{0, [illegible]\} \\ \max\{.4, .3, .5\} & \max\{.4, .2, 0\} & \max\{.4, 0, [illegible]\} & \max\{.4, [illegible]\} \end{bmatrix}$$

$$= \begin{bmatrix} .8 & .3 & .5 & .5 \\ 1 & .9 & .5 & .7 \\ .5 & .4 & .5 & .6 \end{bmatrix}$$

The max. product composition of the fuzzy relation P and Q is

$$\mu_{P \circ Q} = \begin{bmatrix} .3 & .5 & .8 \\ 0 & .7 & 1 \\ .4 & .6 & .5 \end{bmatrix} \cdot \begin{bmatrix} .9 & .5 & .7 & .7 \\ .3 & .2 & 0 & .9 \\ 1 & 0 & .5 & .5 \end{bmatrix}$$

$$= \begin{bmatrix} \max\{.27, .15, .8\} & .15 & .4 & .45 \\ 1 & .14 & .5 & .63 \\ .5 & .2 & .28 & .54 \end{bmatrix}$$

Defn:

Let $R(X, X)$ be a fuzzy relation. Then $R(X, X)$ is called a reflexive if $\mu_R(x, x) = 1$ $\forall x \in X$.

If this is not the case for some $x \in X$, then the relation $R(x, x)$ is called irreflexive. If it is not satisfied for all $x \in X$, then the relation $R(X, X)$ is called anti-reflexive.

A weaker form of reflexivity, referred to as ϵ-reflexivity, is sometimes defined by $\mu_R(x, x) \geq \epsilon$, where $0 < \epsilon < 1$.

Defn: A fuzzy relation is symmetric iff $\mu_R(x, y) = \mu_R(y, x)$ $\forall x, y \in X$. Whenever this equality is not satisfied for some $x, y \in X$, the fuzzy relation is called asymmetric.

$\max\{.3, .5, .5\}$ When $\mu_R(x, y) > 0$ and $\mu_R(y, x) > 0$ implies

$\max\{0, .7, .9\}$ that $x = y$ for all $x, y \in X$, then the fuzzy

$\max\{.4, .6, .5\}$ relation R is called anti symmetric.

Defn: A fuzzy relation $R(X, X)$ is called transitive iff

$\mu_R(x,z) \geq \max_{y \in Y} \min \{ \mu_R(x,y), \mu_R(y,z) \}$

$\forall x, y, z \in X$.

If it is not satisfied for some $x \in X$ then the fuzzy relation $R(X,X)$ is called nontransitive.

If $\mu_R(x,z) < \max_{y \in Y} \min \{ \mu_R(x,y), \mu_R(y,z) \}$ $\forall x, y, z \in X$, then the fuzzy relation R is called anti transitive.

Note:
Another condition for transitive is

$\mu_R(x,z) \geq \max_{y \in Y} \{ \mu_R(x,y) \cdot \mu_R(y,z) \}$

$\forall x, y, z \in X$.

Example: Let X be a collection of cities. The fuzzy relation R is defined on X with respect
5 to " very near ".

soln: Clearly R is reflexive since a city is certainly very near to itself.

If city A is very near to B, then B is certainly very near to A to the same degree.

$\therefore$ R is symmetric.

If city A is very near to B to some degree say 0.7 and city B is very near to C to some degree say 0.8, it is possible that city A is very near to C to a smaller degree say 0.5.

$\therefore$ R is transitive

$\therefore R(X,X)$ is Equivalence Relation.

Transitivity relation :

A binary relation $R(X,X)$ is called equivalence relation if $R(X,X)$ is reflexive, symmetric and transitive.

Algo: A binary relation $R(X,X)$; the last requirement is generalized such that the elements of the transitive closure have the smallest possible membership grade that still allow the first two requirements to be met. the transitive closure $R_T(X,X)$ can be determined by a simple algorithm that consists of the following three steps:

1. $R' = R \cup (R \circ R)$.
2. If $R' \neq R$, make $R = R'$ and go to step 1.
3. stop $R' = R_T$.

Example : Find the transitive max-min closure $R_T(X,X)$ for a fuzzy relation $R(X,X)$ defined by the membership matrix

$$M_R = \begin{bmatrix} .7 & .5 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & .4 & 0 & 0 \\ 0 & 0 & .8 & 0 \end{bmatrix}$$

Soln: Applying steps of the algorithm we obtain

$$M_{R\circ R} = M_R \circ \begin{bmatrix} .7 & .5 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & .4 & 0 & 0 \\ 0 & 0 & .8 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} .7 & .5 & 0 & .5 \\ 0 & 0 & .8 & 0 \\ 0 & 0 & 0 & .4 \\ 0 & .4 & 0 & 0 \end{bmatrix}$$

Now $\mu_{R\cup(R\circ R)} = \begin{bmatrix} .7 & .5 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & .4 & .8 & 0 \\ 0 & 0 & [illegible] & [illegible] \end{bmatrix} \cup \begin{bmatrix} 0 & 0 & .8 & .4 \\ 0 & 0 & 0 & [illegible] \\ 0 & .4 & [illegible] & 0 \\ [illegible] & [illegible] & [illegible] & [illegible] \end{bmatrix}$

$$= \begin{bmatrix} .7 & .5 & 0 & 0.5 \\ 0 & 0 & .8 & 1 \\ 0 & .4 & 0 & .4 \\ 0 & .4 & .8 & 0 \end{bmatrix}$$

$$= \mu_{R'}$$

Since $R' \neq R$, we take as a new relation R and repeating the previous procedure, we obtain

$$\mu_{R\circ R} = \begin{bmatrix} .7 & .5 & 0 & .5 \\ 0 & 0 & .8 & 1 \\ 0 & .4 & 0 & .4 \\ 0 & .4 & .8 & 0 \end{bmatrix} \begin{bmatrix} .7 & .5 & 0 & .5 \\ 0 & 0 & .8 & 1 \\ 0 & .4 & 0 & .4 \\ 0 & .4 & .8 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} .7 & .5 & .5 & .5 \\ 0 & .4 & .8 & .4 \\ 0 & .4 & .4 & .4 \\ 0 & .4 & .4 & .4 \end{bmatrix}$$

$$\mu_{R\cup(R\circ R)} = \begin{bmatrix} .7 & .5 & 0 & .5 \\ 0 & 0 & .8 & 1 \\ 0 & .4 & 0 & .4 \\ 0 & .4 & .8 & 0 \end{bmatrix} \cup \begin{bmatrix} .7 & .5 & .5 & .5 \\ 0 & .4 & .8 & .4 \\ 0 & .4 & .4 & .4 \\ 0 & .4 & .4 & .4 \end{bmatrix}$$

$$= \begin{bmatrix} .7 & .5 & .5 & .5 \\ 0 & .4 & .8 & 1 \\ 0 & .4 & .4 & .4 \\ 0 & .4 & .8 & .4 \end{bmatrix} = \mu_{R'}$$

$R' \neq R$, we take as a new relation R, and repeating the previous procedure, we obtain

$$\mu_{R\circ R} = \begin{bmatrix} .7 & .5 & .5 & .5 \\ 0 & .4 & .8 & 1 \\ 0 & .4 & .4 & .4 \\ 0 & .4 & .8 & .4 \end{bmatrix} \begin{bmatrix} .7 & .5 & .5 & .5 \\ 0 & .4 & .8 & 1 \\ 0 & .4 & .4 & .4 \\ 0 & .4 & .8 & .4 \end{bmatrix}$$

$$\begin{bmatrix} \cdot 4 & \cdot 5 & \cdot 5 & \cdot 5 \\ 0 & \cdot 4 & \cdot 8 & \cdot 4 \\ 0 & \cdot 4 & \cdot 4 & \cdot 4 \\ 0 & \cdot 4 & \cdot 4 & \cdot 4 \end{bmatrix}$$

$$\therefore \mu_{R\cup(R\circ R)} = \begin{bmatrix} \cdot 4 & \cdot 5 & \cdot 5 & \cdot 5 \\ 0 & \cdot 4 & \cdot 8 & \cdot 4 \\ 0 & \cdot 4 & \cdot 4 & \cdot 4 \\ 0 & \cdot 4 & \cdot 8 & \cdot 4 \end{bmatrix} \begin{bmatrix} \cdot 4 & \cdot 5 & \cdot 5 & \cdot 5 \\ 0 & \cdot 4 & \cdot 8 & \cdot 4 \\ 0 & \cdot 4 & \cdot 4 & \cdot 4 \\ 0 & \cdot 4 & \cdot 4 & \cdot 4 \end{bmatrix}$$

$$= \begin{bmatrix} \cdot 4 & \cdot 5 & \cdot 5 & \cdot 5 \\ 0 & \cdot 4 & \cdot 8 & \cdot 4 \\ 0 & \cdot 4 & \cdot 4 & \cdot 4 \\ 0 & \cdot 4 & \cdot 8 & \cdot 4 \end{bmatrix} = \mu_{R'}$$

$\therefore \mu_R = \mu_{R'}$ i.e. $R' = R$

The transitive closure R_T whose membership matrix is $\mu_{R_T} = \begin{bmatrix} \cdot 4 & \cdot 5 & \cdot 5 & \cdot 5 \\ 0 & \cdot 4 & \cdot 8 & 1 \\ 0 & \cdot 4 & \cdot 4 & \cdot 4 \\ 0 & \cdot 4 & \cdot 8 & \cdot 4 \end{bmatrix}$

8

Defn: The equivalence class A_x with respect to x of the equivalence relation $R(X,X)$ is defined as $A_x = \{ y / (x,y) \in R(X,X) \}$.

Note: The collection of all equivalence classes of $R(X,X)$ is dinoted by X/R (μ_{X}) which is a partition of X.

Example: Let $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, let (x, y) have the same $R(X,X)$ and x and y have the same remainder when divided by 3. i) show that the Relation R is equivalence relation.

ii) Find the partition of X or find X/R.

Example:

The fuzzy relation $R(X, X)$ represented by the Membership matrix

$$
\begin{array}{c|c}
 & \begin{array}{ccccccc} a & b & c & d & e & f & g \end{array} \\
\begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} &
\begin{bmatrix}
1 & .8 & 0 & .4 & 0 & 0 & 0 \\
.8 & 1 & 0 & .4 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & .9 & .9 & .5 \\
.4 & .4 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & .9 & 0 & 1 & .9 & .5 \\
0 & 0 & .9 & 0 & .9 & 1 & .5 \\
0 & 0 & .5 & 0 & .5 & .5 & 1
\end{bmatrix}
\end{array}
$$

praise the Lord

i) Show that the fuzzy relation $R(X, X)$ is equivalence fuzzy relation (similarity)

ii) Write the partition tree of the fuzzy relation $R(X, X)$

soln: i) clearly $R(X, X)$ is reflexive since $\mu_R(x, x) = 1 \;\forall x \in X$

Given matrix is symmetric

$R(X, X)$ is symmetric

From the membership matrix of the fuzzy relation $R(X, X)$ is transitive.

$R(X, X)$ is similarity relation

ii)

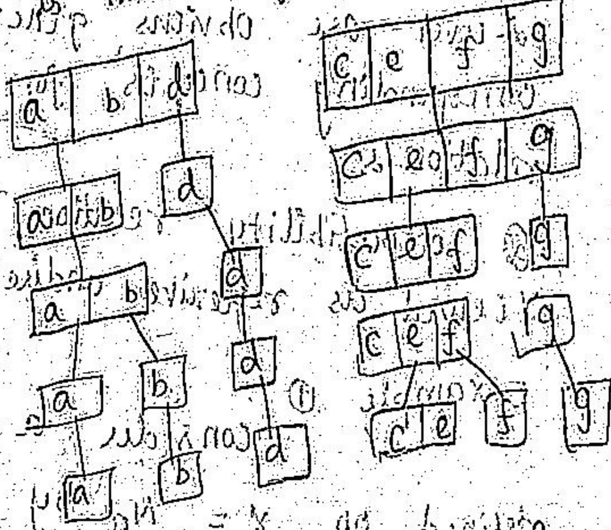
$\alpha = .4$

$.5$

$\alpha = .8$

$\alpha = .9$

$\alpha = 1$



Def: A binary relation $R(X, X)$ is called a reflexive and symmetric compatibility relation or tolerance relation.

When $R(x,x)$ is a reflexive and symmetric fuzzy relation, then $R(x,x)$ is called a proximity relation.

Defn : Given a crisp compatibility relation $R(x,x)$ a compatibility class is a subset A of X such that $\langle x, y \rangle \in R$ $\forall x, y \in A$.

Defn : A maximal compatibility class or maximal compatible is a compatibility class that is not properly contained within any other compatibility class.

Defn : The family consisting of all the maximal compatibles induced by R on X is called a complete cover of X with respect to R.

Defn : An α- compatibility class is a subset A of X such that $R(x,y) \geq \alpha$ $\forall x, y \in A$

Note ① : Maximal α- compatibles and complete α-cover are obvious generalization of the corresponding concepts for crisp compatibility relations.

② compatibility relations are after conveniently viewed as reflexive undirected graphs.

Example ① : Consider a fuzzy relation $R(x,x)$ defined on $X = N_9$ by the following membership matrix

$$
\begin{array}{c|ccccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
\hline
1 & 1 & .8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & .8 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 1 & 1 & .8 & 0 & 0 & 0 & 0 \\
4 & 0 & 0 & 1 & 1 & .8 & .7 & .5 & 0 & 0 \\
5 & 0 & 0 & .8 & .8 & 1 & .7 & .4 & .7 & 0 \\
6 & 0 & 0 & 0 & .7 & .7 & 1 & .4 & 0 & 0 \\
7 & 0 & 0 & 0 & .5 & .4 & .4 & 1 & 0 & 0 \\
8 & 0 & 0 & 0 & 0 & .7 & 0 & 0 & 1 & 0 \\
9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}
$$

The relation is compatibility relation.

i) verify the relation is compatibility relation.

ii) Draw the graph.

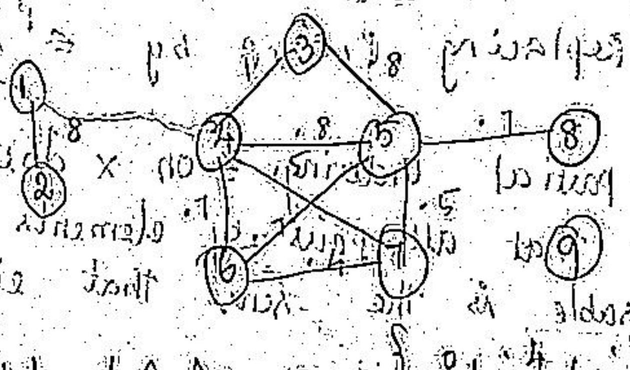
iii) Find the complete α-cover

soln: i) All the entries on the main diagonal are equal to 1 ⇒ R is reflexive

Given membership matrix is symmetric ⇒ $R(x, y)$ is symmetric

⇒ R is compatibility relation

(ii)



iii) Let $\alpha \in \Lambda_R = \{0, .4, .5, .7, .8, 1\}$

Complete α-covers for $\alpha > 0$ is $\alpha = 1$

$\alpha = .4$: [1 2] [3 4 5] [4 5 6 7] [5 8] [9]

$\alpha = .5$: [1 2] [3 4 5] [4 5 6] [4 5 7] [5 8] [9]

$\alpha = .7$: [1 2] [3 4 5] [4 5 6] [4 7] [5 8] [9]

$\alpha = .8$: [1 2] [3 4 5] [6] [7] [8 5] [9]

$\alpha = 1$: [1] [2] [3] [4] [5] [6] [7] [8] [9]

UNIT-IV

Fuzzy measures:

Definition:

A fuzzy measure is defined by a function $g : \rho(x) - [0.1]$ which assigns to each crisp subset of X a number in the unit interval [0.1].

Axioms of fuzzy measures:

1. Boundary conditions $g(\emptyset) = 0$ and $g(x) = 1$
2. Monotonicity: For every $A, B \in \rho(x)$. if $A \leq B$ then $g(A) \leq g(B)$
3. Continuity: For every sequence $A_i \in \rho(x) / i \in N$ of subsets of X, if either $A_1 \subseteq A_2 \subseteq \cdots \ldots\ldots$ or $A_1 \supseteq A_2 \supseteq \cdots \ldots\ldots$ then $\lim_{i \to \infty} g(A_i) = g(\lim_{i \to \infty}(A_i))$

Definition:

A fuzzy measure is often defined more generally as a function $g : \beta - [0.1]$, where $\beta \subseteq \rho(x)$ is a family of subsets of X such that

i. $\emptyset, X \in \beta$
ii. If $A \in \beta$, then $\bar{A} \in \beta$
iii. β is closed under the operation of set union, that is if $A \in \beta, B \in \beta$, then also $A \cup B \in \beta$

The set β is usually called a Borel field or σ –field

Note:

i. Since $A \cup B \supseteq A$ and $A \cup B \supseteq B$, we have $\max\{g(A), g(B)\} \leq g(A \cup B)$ due to the requi monotonicity of the function g.
ii. Since $A \cup B \subseteq A$ and $A \cup B \subseteq B$, we have $\max\{g(A), g(B)\} \leq \min g(A \cup B)$

Definition:

A belief measure is a function $Bel : \rho(x) - [0.1]$ that satisfies

i. $Bel(\emptyset) = 0$ and $Bel(X) = 1$
ii. For every sequence $A_i \in \rho(x) / i \in N$ of subsets of X, if $A_1 \subseteq A_2 \subseteq \cdots \ldots\ldots$ or $A_1 \supseteq A_2 \supseteq \cdots \ldots\ldots$ then $\lim_{i \to \infty} Bel(A_i) = Bel(\lim_{i \to \infty}(A_i))$
iii. $Bel(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) \geq \sum_i Bel(A_i) - \sum_i Bel(A_i \cap A_j) + \cdots +$ $(-1)^{n+1} Bel(A_1 \cap A_2 \cap \ldots\ldots \cap A_n)$

Definition:

Let $R_\alpha = \{x / x \in X, \mu_{\tilde{R}}(x) \geq \alpha\}$ be the α- level sets of the solution space and $N(\alpha) = \{x / x \in R_\alpha, f(x) = \sup_{x' \in R_\alpha} f(x')\}$, the set of optimal solutions for each α- level set. The function set "decision" is then defined by the membership function $\mu_{opt}(x) = \begin{cases} \sup_{x \in N_\alpha} \alpha, & if\ x \in \cup_{\alpha > 0} N(\alpha) \\ 0, & else \end{cases}$

The fuzzy set "optimal values of the objective function" has the membership function

$\mu_f(r) = \begin{cases} \sup_{x \in f^{-1}(r)} \mu_{opt}(x), & if\ x \in R_1 \wedge f^{-1}(r) \neq 0 \\ 0, & else \end{cases}$

$f(x)$ is the objective function with functional values r.

The determination of the r's and $\mu_{opt}(x)$ can be obtained by parametric programming.

For each α and LP of the following kind would have to be solved.

$maximize\ f(x)$

Subject to constraint $\alpha \leq \mu_i(x),\ i = 1, 2, \ldots, m,\ x \in X$

Note:

1. For n=2, we have

$Bel\,(A_1 \cup A_2) \geq Bel(A_1) + Bel(A_2) \ldots - Bel\,(A_1 \cap A_2)$

For n=3, we have

$Bel\,(A_1 \cup A_2 \cup A_3) \geq Bel(A_1) + Bel(A_2) + Bel(A_3) \ldots - Bel\,(A_1 \cap A_2) - Bel\,(A_2 \cap A_3)$

$-Bel\,(A_2 \cap A_3) + Bel\,(A_1 \cap A_2 \cap A_3)$

2. Let $A \subseteq B$ and let $C = B - A$. Then $A \cup C = B$ and $A \cup C = \emptyset$
Now applying A and C to note (1), we obtain

$Bel\,(A \cup C) = Bel(B) \geq Bel(A) + Bel(C) - Bel(A \cap C)$ and $A \cap C = \emptyset$

$\therefore Bel\,(\emptyset) = 0$, we have $Bel\,(A \cup C) \geq Bel(A) + Bel(C)$

Consequently $Bel(B) \geq Bel(A)$
3. Let $A_1 = A$ and $A_2 = \bar{A}$. Then $Bel\,(A \cup \bar{A}) \geq Bel(A) + Bel(\bar{A}) - Bel(A \cap \bar{A})$
$1 \geq Bel(A) + Bel(\bar{A})$

Definition:
A plausibility measure is a function $pl : p(x) \to [0,1]$ that satisfies

i. $pl(\emptyset) = 0$ and $pl(X) = 1$
ii. For every sequence $A_i \in p(x) / i \in N$ of subsets of X, if either $A_1 \subseteq A_2 \subseteq \ldots\ldots$ or $A_1 \supseteq A_2 \supseteq \ldots\ldots$ then $\lim_{i\to\infty} pl(A_i) = pl(\lim_{i\to\infty}(A_i))$
iii. $pl\,(A_1 \cup A_2 \cup \ldots \ldots \cup A_n) \geq \sum_i pl(A_i) - \sum_{i<j} pl(A_i \cap A_j) + \cdots + (-1)^{n-1} pl\,(A_1 \cup A_2 \cup \ldots \ldots \cup A_n)$ for every $n \in N$ and every collection of subsets of X.

Note:

1. The relation between pl and Bel is defined by $pl(A) = 1 - Bel\,(\bar{A})$ for all $A \in p(x)$ (or) $Bel(A) = 1 - pl\,(\bar{A})$
2. Let n = 2, $A_1 = A$ and $A_2 = \bar{A}$. Then $pl(A \cap \bar{A}) \leq pl(A) + pl(\bar{A}) - pl(A \cup \bar{A})$
$0 \leq pl(A) + pl(\bar{A}) - 1$
$1 \leq pl(A) + pl(\bar{A})$
3. We know that
$Bel\,(A_1 \cup A_2 \cup A_3) \geq Bel(A_1) + Bel(A_2) + Bel(A_3) \ldots - Bel\,(A_1 \cap A_2) - Bel\,(A_2 \cap A_3)$

$-Bel\,(A_2 \cap A_3) + Bel\,(A_1 \cap A_2 \cap A_3)$

$1 - pl\,(\overline{(A_1 \cup A_2 \cup A_3)}) \geq 1 - pl(\bar{A}_1) + 1 - pl(\bar{A}_2) + 1 - pl(\bar{A}_3) \ldots - [1 - pl\,(A_1 \cap A_2)]$
$-[1 - pl\,(A_1 \cap A_3)] - [1 - pl\,(A_2 \cap A_3)] + [1 - pl\,(A_1 \cap A_2 \cap A_3)]$

Definition:

A basic assignment function m is defined as $m : p(x) \to [0,1]$ such that $m(\emptyset) = 0$ and $\sum_{A \in p(x)} m(A) = 1$

Result:

1. $Bel(A) = \sum_{B \subseteq A} m(B)$ and $pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$
2. $pl(A) \geq Bel(A)$

Definition:

Every set $A \in p(x)$ for which $m(A) > 0$ is usually called a focal element of m. As this name suggests, focal elements are subsets of X on which the available evidence focuses. When X is finit3e, m can be fully characterized by a list of its focal elements A with the corresponding values $m(A)$. The pair (f, m), where f and m denote a set of focal elements and the associated basic assignment respectively is often called a body of evidence.

Definition:

Total ignorance is expressed in terms of the basic assignment by $m(X) = 1$ and $m(A) = 0$ for all $A \neq X$.

Definition:

A basic assignment m is said to be a simple support function focused at A of $m(A) = S, m(X) = 1 - S$ and $m(B) = 0$ for all others sets B in $p(X)$.

Note: where denotes the cardi...

Result

1. Let two basic assignments m_1 and m_2 on same power set $p(x)$. Then they must be appointedly combined to obtain a joint basic assignment $m_{1,2}$ which is defined as $m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_1(B) \cdot m_2(C)}{1-K}, \forall A \neq \emptyset$, where $K = \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C), m_{1,2}(\emptyset) = 0$. This is called on Dempster''s rule of combination.

2. The sum of products $m_1(B) \cdot m_2(C)$ for all focal elements B of m_1 and all focal elements C of m_2 such that $B \cap A \neq \emptyset$ is equal to $1 - K$.

Problem 1:

Let R, D and C denote subsets of our universal set X. The set of all paintings that contain the set of all painting by Raphael, the set of all paintings by disciples of Raphael and the set of all counterfeits of Raphael''s and the set of all counterfeits of Raphael''s paintings respectively.

Now that two experts performed careful examination of the painting and subsequently provided us with basic assignments m_1 and m_2 specified in table:

Focal elements		R	D	C	$R \cup D$	$R \cup C$	$D \cup C$	$R \cup D \cup C$
Expert I	m_1	0.05	0	0.05	0.15	0.1	0.05	0.6
Expert II	m_2	0.15	0	0.05	0.05	0.2	0.05	0.5

Find the belief and plausibility measures and calculate the joint basic assignment $m_{1,2}$, $Bel_{1,2}$.

Solution:

We know that $Bel(A) = \sum_{B \subseteq A} m(B)$ and $pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$

Hence

Focal elements	Expert I			Expert II			Combined evidence		
	m_1	Bel_1	pl_1	m_2	Bel_2	pl_2	$m_{1,2}$	$Bel_{1,2}$	$pl_{1,2}$
R	0.05	0.05	0.9	0.15	0.15	0.75	0.21	0.21	0.84
D	0	0	0.8	0	0	0.6	0.01	0.01	0.5
C	0.05	0.05	0.8	0.05	0.05	0.8	0.09	0.09	0.66

$R \cup D$	0.15	0.2	0.95	0.05	0.2	0.95	0.12	0.34	0.91
$R \cup C$	0.1	0.2	1	0.2	0.4	1	0.2	0.5	0.99
$D \cup C$	0.05	0.1	0.95	0.05	0.1	0.85		0.16	0.79
$R \cup D \cup C$	0.6	1	1	0.5	1	1			

$$m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_1(B) \cdot m_2(C)}{1-k},\ \text{where } k = \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C),\ m_{1,2}(\emptyset)$$

Since, the normalization factor

$$k = m_1(R) \cdot m_2(D) + m_1(R) \cdot m_2(C) + m_1(R) \cdot m_2(D \cup C) + m_1(D) \cdot m_2(R) + m_1(D) \cdot m_2(C)$$

$$+ m_1(D) \cdot m_2(R \cup C) + m_1(C) \cdot m_2(R) + m_1(C) \cdot m_2(D) + m_1(C) \cdot m_2(R \cup D)$$

$$+ m_1(R \cup D) \cdot m_2(C) + m_1(R \cup C) \cdot m_2(D) + m_1(D \cup C) \cdot m_2(R)$$

$$= (0.05)(0) + (0.05)(0.05) + (0.05)(0.05) + (0)(0.15) + (0)(0.05) + (0)(0.2)$$

$$+ (0.05)(0.15) + (0.05)(0) + (0.05)(0.05) + (0.15)(0.05) + (0.1)(0) + (0.05)(0.15)$$

$$= 0.03$$

$$1 - k = 1 - 0.03 = 0.97$$

Now,

$$m_{1,2}(R) = \left[\begin{array}{c} m_1(R) \cdot m_2(R) + m_1(R) \cdot m_2(R \cup D) + m_1(R) \cdot m_2(R \cup C) + m_1(R) \cdot m_2(R \cup D \cup C) \\ + m_1(R \cup D) \cdot m_2(R) + m_1(R \cup D) \cdot m_2(R \cup C) + m_1(R \cup C) \cdot m_2(R) + m_1(R \cup C) \cdot m_2(R) \\ + m_1(R \cup C) \cdot m_2(R \cup D) + m_1(R \cup D \cup C) \cdot m_2(R) \end{array}\right] \Big/ 0.97$$

$$= 0.21$$

Proceeding this we get $m_{1,2}, Bel_{1,2}, pl_{1,2}$

Problem 2:

Let $X = \{a, b, c, d\}$. Given the belief measure $Bel(\{b\}) = 0.1, Bel(\{a, b\}) = 0.2, Bel(\{b, c\}) = 0.3, Bel(\{b, d\}) = 0.1, Bel(\{a, b, c\}) = 0.4, Bel(\{a, b, d\}) = 0.2, Bel(\{b, c, d\}) = 0.6, Bel(X) = 1$, determine the corresponding basic assignment.

Solution:

Bel	m	Bel	m	Bel	m
$0 = Bel(\{a\})$	0	$0.2 = Bel(\{a,b\})$	0.1	$0 = Bel(\{c,d\})$	0
$0.1 = Bel(\{b\})$	0.1	$0 = Bel(\{a,c\})$	0	$0.4 = Bel(\{a,b,c\})$	0
$0 = Bel(\{c\})$	0	$0 = Bel(\{a,d\})$	0	$0.2 = Bel(\{a,b,d\})$	0
$0 = Bel(\{d\})$	0	$0.3 = Bel(\{b,c\})$	0.2	$0.6 = Bel(\{b,c,d\})$	0.3
				$0 = Bel(\{a,c,d\})$	0
				$1 = Bel(\{a,b,c,d\})$	0.3

Since $m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Bel\ B$

$$m(\{a,b\}) = Bel(\{a,b\}) - Bel(\{b\}) - Bel(\{a\})$$

$$= 0.2 - 0.1 - 0 = 0.1$$

$m(\{a,c\}) = Bel(\{a,c\}) - Bel(\{a\}) - Bel(\{c\}) = 0$ and so on.

$m(\{a,b,c\}) = Bel(\{a,b,c\}) - Bel(\{a,b\}) - Bel(\{a,c\}) - Bel(\{b,c\}) + Bel(\{a\}) + Bel(\{b\}) + Bel(\{c\})$

$= 0.4 - 0.2 - 0 - 0.3 + 0 + 0.1 + 0 = 0$

$m(\{a,b,d\}) = Bel(\{a,b,d\}) - Bel(\{a,b\}) - Bel(\{a,d\}) - Bel(\{b,d\}) + Bel(\{a\}) + Bel(\{b\}) + Bel(\{c\})$

$= 0.2 - 0.2 - 0 - 0.1 + 0 + 0 + 0.1 = 0$

$m(\{b,c,d\}) = Bel(\{b,c,d\}) - Bel(\{b,c\}) - Bel(\{b,d\}) - Bel(\{c,d\}) + Bel(\{d\}) + Bel(\{b\}) + Bel(\{c\})$

$= 0.6 - 0.3 - 0.1 - 0 + 0.1 + 0 + 0 = 0.3$

$m(\{a,c,d\}) = Bel(\{a,c,d\}) - Bel(\{a,c\}) - Bel(\{a,d\}) - Bel(\{c,d\}) + Bel(\{a\}) + Bel(\{a\}) + Bel(\{c\})$

$= 0 - 0 - 0 - 0 + 0 + 0 + 0 = 0$

$m(\{a,b,c,d\}) = Bel(\{a,b,c,d\}) - Bel\{a,b,c\} - Bel\{a,b,d\} - Bel\{a,c,d\} - Bel\{b,c,d\}$

$+ Bel\{a,b\} + Bel(\{a,c\}) + Bel(\{a,d\}) + Bel(\{b,c\}) + Bel(\{b,d\})$

$+ Bel(\{c,d\}) - Bel(\{a\}) - Bel(\{a\}) - Bel(\{c\}) - Bel(\{d\})$

$= 1 - 0.6 - 0.2 - 0.4 + 0.1 + 0.3 + 0.2 - 0.1$

$= 1.6 - 1.3 = 0.3$

Definition:

In a Cartesian product $Z = X \times Y$. The basic assignment function m is defined as $m : P(X \times Y) \to [0,1]$ each focal element of m is in this case a binary relation R on $X \times Y$.

Definition:

In a crisp set the relation R is a subset of $X \times Y$. Then R_x is a projection of R on X is defined as

$R_x = \{x \in X/(x,y) \in R \; for \; some \; y \in Y\}$ and $R_y = \{y \in Y/(x,y) \in R \; for \; some \; x \in X\}$ defines the projection of R on Y.

Definition:

We define the projection m_x of m on X by the formula $m_x(A) = \sum_{R/A = R_x} m(R), for \; all \; A \in P(X)$.

Note:

To calculate $m_x(A)$ according to this formula, we add the values of $m(R)$ for all focal elements R

Definition:

We define the projection m_y of m on Y by the formula $m_y(B) = \sum_{R/B = R_y} m(R), for \; all \; B \in P(Y)$.

Result:

Let m_x and m_y be called marginal basic assignments and let (f_x, m_x) and (f_y, m_y) be the associated marginal bodies of evidence

Definition:

Two marginal bodies of evidence (f_x, m_x) and (f_y, m_y) are said to be non-iteractive iff for all $A \in f_x$ and $B \in f_y$, $m(A \times B) = m_x(A) \cdot m_y(B)$ and $m(R) = 0$ for all $R \neq A \times B$.

Problem 1:

Consider the body of evidence given in the table. Focal elements are subsets of the Cartesian product $X \times Y$, where $X = \{1,2,3\}$ and $Y = \{a, b, c\}$. They are defined in the table by their characteristic functions. To emphasize that each focal element is, infact abinary relation $X \times Y$, they are labeled $R_1, R_2, \ldots, R_{12}$.

R_i	X×Y									$m(R_i)$
	1a=(1,a)	1b	1c	2a	2b	2c	3a	3b	3c	
R_1	0	0	0	0	1	1	0	1	1	0.0625
R_2	0	0	0	1	0	0	1	0	0	0.0625
R_3	0	0	0	1	1	1	1	1	1	0.125
R_4	0	1	1	0	0	0	0	1	1	0.0375
R_5	0	1	1	0	1	1	0	0	0	0.075
R_6	0	1	1	0	1	1	0	1	1	0.075
R_7	1	0	0	0	0	0	1	0	0	0.0375
R_8	1	0	0	1	0	0	0	0	0	0.075
R_9	1	0	0	1	0	0	1	0	0	0.075
R_{10}	1	1	1	0	0	0	1	1	1	0.075
R_{11}	1	1	1	1	1	1	0	0	0	0.15
R_{12}	1	1	1	1	1	1	1	1	1	0.15

i. Show that the marginal bodies of evidence.

ii. The marginal bodies of evidence are non-interavtive

Solution:

(i) Consider,

$m_x(\{1\}) = m_x(\{2\}) = m_x(\{3\}) = 0$ Since

$m_x(\{1,2\}) = m(R_5) + m(R_8) + m(R_{11})$

$= 0.075 + 0.075 + 0.15 = 0.3$

$m_x(\{1,3\}) = m(R_4) + m(R_7) + m(R_{10})$

$= 0.0375 + 0.0375 + 0.075 = 0.15$

$m_x(\{2,3\}) = m(R_1) + m(R_2) + m(R_3)$

$= 0.0625 + 0.0625 + 0.125 = 0.25$

$m_x(\{1,2,3\}) = 0.3$

$m_y(\{a\}) = 0.25$

$m_y(\{b\}) = 0$

$m_y(\{c\}) = 0$

$m_y(\{a,c\}) = 0$

$m_y(\{b,c\}) = 0.25$

∴The marginal bodies of evidence

X			$m_X(A)$
1	2	3	
0	1	1	0.25
1	0	1	0.15
1	1	0	0.3
1	1	1	0.3

Y			$m_Y(B)$
a	b	c	
0	1	1	0.25
1	0	0	0.25
1	1	1	0.5

(ii) Consider,

$m(R_1) = m(\{2,3\} \times \{b, c\})$

$= m_X(\{2,3\}) \cdot m_Y(\{b, c\})$

$= 0.25 \times 0.25 = 0.0625$

$m(R_2) = m(\{2,3\} \times \{a\})$

$= m_X(\{2,3\}) \cdot m_Y(\{a\})$

$= 0.25 \times 0.25 = 0.0625$

$m(R_3) = m(\{2,3\} \times \{a, b, c\})$

$= m_X(\{2,3\}) \cdot m_Y(\{a, b, c\})$

$= 0.25 \times 0.5 = 0.125$

Similarly,

$m(R_4) = 0.0375 \qquad m(R_8) = 0.075$

$m(R_5) = 0.075 \qquad m(R_9) = 0.075$

$m(R_6) = 0.075 \qquad m(R_{10}) = 0.075$

$m(R_7) = 0.075 \qquad m(R_{11}) = 0.15$

∴The marginal bodies are of evidence are non-interactive.

Definition:

The probability measures is a function $Bel : P(X) \to [0,1]$ which satisfies the following conditions:

i.

ii. For every sequence $A_i \in P(X) / i \in N$ of subsets of X, if either $A_1 \subseteq A_2 \subseteq \dots\dots$ or $A_1 \supseteq A_2 \supseteq \dots\dots$ then $\lim_{i\to\infty} Bel(A_i) = Bel(\lim_{i\to\infty}(A_i))$

iii. $Bel(A \cup B) = Bel(A) + Bel(B)$, whenever $A \cap B = \emptyset$. It is also called Bayesian belief measure.

Theorem:

A belief measure Bel on a finite power set $P(X)$ is a probability measure if and only if its basic assignment m is given by $m(\{x\}) = Bel(\{x\})$ and $m(A) = 0$ for all subsets of X that are not singletons.

Proof:

Assume that Bel is a probability measure. For the empty set $\emptyset$, the theorem trivially holds, since $m(\emptyset) = 0$ by definition of m. Let $A = \emptyset$ and assume $A = \{x_1, x_2, \dots\dots, x_n\}$.

Now, $Bel(A) = Bel(\{x_1\}) + Bel(\{x_2, x_3, \dots\dots, x_n\})$,

by definition $= Bel(\{x_1\}) + Bel(\{x_2\}) + Bel(\{x_3, x_4, \dots\dots, x_n\})$,

$= Bel(\{x_1\}) + Bel(\{x_2\}) + \dots\dots + Bel(\{x_n\})$

$= m(\{x_1\}) + m(\{x_2\}) + \dots\dots + m(\{x_n\})$

$$Bel(A) = \sum_{i=1}^{n} m(\{x_i\})$$

Hence Bel is defined in terms of a basic assignment that focuses only on singletons.

Conversely, assume that a basic assignment m is given such that $\sum_{x \in X} m(\{x\}) = 1$

Then for any sets $A, B \in P(X)$ such that $A \cap B = \emptyset$,

We have, $Bel(A) + Bel(B) = \sum_{x \in A} m(\{x\}) + \sum_{x \in B} m(\{x\})$

$= \sum_{x \in A \cup B} m(\{x\}) = Bel(A \cup B)$

$Bel(A \cup B) = Bel(A) + Bel(B)$

$\therefore Bel$ is a probability measure.

Definition:

The probability measures on finite sets are thus fully represented by a function $P: X \to [0,1]$ such that $P(x) = m(\{x\})$. This function is usually called a probability distribution function.

Result:

1. Let $P = \langle P(x)/x \in X \rangle$ be referred to as a probability distribution on X.
2. Here $Bel(A) = pl(A) = \sum_{x \in A} m(\{x\}), \forall A \in P(X)$
3. $Bel(A) = pl(A) = \sum_{x \in A} P(x)$
4. Let us denote probability measures by P. Then $P(A) = \sum_{x \in A} P(x), \forall A \in P(X)$

Definition:

When a probability distribution P is defined on the cartesian product $X \times Y$, it is called a joint probability distribution projections P_X and P_Y of P on X and Y respectively are called marginal probability distributions, they are defined by the formulas

$P_X = \sum_{y \in Y} P(x, y); for\ each\ x \in X$ and

$$P_Y = \sum_{x \in X} P(x, y); for\ each\ y \in Y.$$

Definition:

Sets X and Y are called non-interactive with respect to P if $P(x, y) = P_X(x).P_Y(y), \forall x \in X, y \in Y$

Definition:

Two conditional probability distributions $P_{X/Y}$ and $P_{Y/X}$ are defined in terms of a joint distribution P by the formulas,

$P_{X/Y}(x/y) = \frac{P(x, y)}{P_Y(y)}$,

$$P_{Y/X}(y/x) = \frac{P(x,y)}{P_X(x)}, \forall x \in X, y \in Y$$

Definition:

Set X is called independent of Y if $P_{X/Y}(x/y) = P_X(x), \forall x \in X, y \in Y$

Set Y is called independent of X if $P_{Y/X}(y/x) = P_Y(y), \forall x \in X, y \in Y$

Note:

The joint probability distribution is defined as $P_{X/Y}(x/y) \cdot P_Y(y) = P_{Y/X}(y/x) \cdot P_X(x)$

Result:

When it is required that the focal elements of a body evidence (f, m) be nested, the associated belief and plausibility measures are called consonant.

Theorem:

Given a consonant body of evidence (f, m), the associated consonant belief and plausibility measures are possess the following properties:

i. $Bel(A \cap B) = \min[Bel(A), Bel(B)]$
ii. $pl(A \cup B) = \max[pl(A), pl(B)], \forall A, B \in P(X)$

Proof:

i. Since the focal elements in f are nested, they may be linearly ordered by the subset relationship.
let $f = A_1, A_2, \ldots, A_n$ and assume that $A_i \subset A_j$ whenever $i < j$

Consider now arbitrary subsets A and B of X. Let i_1 be that largest integer I such that $A_i \subseteq B$.

Then $A_i \subseteq A$ and $A_i \subseteq B$ iff $i \leq i_1, i \leq i_2$ respectively.

$\therefore A_i \cap A_i \subseteq A \cap B \Rightarrow A_i \subseteq A \cap B$ iff $i \leq \min(i_1, i_2)$

Hence $\quad Bel(A \cap B) = \sum_{i=1}^{\min(i_1, i_2)} m(A_i)$

$$= \min\left\{\sum_{i=1}^{\min(i_1)} m(A_i), \sum_{i=1}^{\min(i_2)} m(A_i)\right\}$$

$Bel(A \cap B) = \min[Bel(A), Bel(B)]$

ii. $\quad pl(A \cup B) = 1 - Bel(\overline{A \cup B})$

$= 1 - Bel(\bar{A} \cap \bar{B})$

$= 1 - \min[Bel(\bar{A}), Bel(\bar{B})]$

$= \max(1 - Bel(\bar{A}), 1 - Bel(\bar{B}))$

$pl(A \cup B) = \max[pl(A), pl(B)], \forall A, B \in P(X)$

Definition:

Consonant belief and plausibility measures are usually referred to as necessity measures and possibility measures respectively. Let η and π denote a necessity measure and a possibility measure on $P(X)$, respectively. Then

$$\eta(A \cap B) = \min\{\eta(A), \eta(B)\} \text{ and}$$

$$\pi(A \cap B) = \min\{\pi(A), \pi(B)\}, \forall A, B \in P(X)$$

Result: $\eta(A) = 1 - \pi(\bar{A}), \forall A \in P(X)$

Theorem:

Every possibility measure π on $P(X)$ can be uniquely determined by a possibility distribution function $r : X \to [0,1]$ and $\pi(A) = \max_{x \in A} r(x)$, for each $A \in P(X)$

Proof:

We prove the theorem by induction on the cardinality of set A.

Let $|A| = 1$. Then $A = \{x\}$, where $x \in X$ and $\pi(A) = \max_{x \in A} r(x)$ is trivially satisfied.

Assume now that $\pi(A) = \max_{x \in A} r(x)$, is satisfied for $|A| = n - 1$ and let $A = \{x_1, x_2, \ldots\ldots, x_n\}$.

Now, $\pi(A) = \max\{\pi(\{x_1, x_2, \ldots\ldots, x_{n-1}\}), \pi(x_n)\}$

$= \max\{\max\{\pi(x_1), \pi(x_2), \ldots\ldots, \pi(x_{n-1})\}, \pi(x_n)\}$

$= \max\{\pi(x_1), \pi(x_2), \ldots\ldots, \pi(x_n)\}$

$= \max_{x \in A} r(x)$

NCERTAINTY-BASED INFORMATION

1 INFORMATION AND UNCERTAINTY

be concept of information, as a subject of this chapter, is intimately connected with the oncept of uncertainty. The most fundamental aspect of this connection is that uncertainty nvolved in any problem-solving situation is a result of some information deficiency. nformation (pertaining to the model within which the situation is conceptualized) may be ncomplete, imprecise, fragmentary, not fully reliable, vague, contradictory, or deficient in ome other way. In general, these various information deficiencies may result in different ypes of uncertainty.

Assume that we can measure the amount of uncertainty involved in a problem-solving situation conceptualized in a particular mathematical theory. Assume further that the amount of uncertainty can be reduced by obtaining relevant information as a result of some action (finding a relevant new fact, designing a relevant experiment and observing the experimental outcome, receiving a requested message, or discovering a relevant historical record). Then, the amount of information obtained by the action may be measured by the reduction of uncertainty that results from the action.

Information measured solely by the reduction of uncertainty does not capture the rich notion of information in human communication and cognition. It is not explicitly concerned with semantic and pragmatic aspects of information viewed in the broader sense. This does not mean, however, that information viewed in terms of uncertainty reduction ignores these aspects. It does not ignore them, but they are assumed to be fixed in each particular application. Furthermore, the notion of information as uncertainty reduction is sufficiently rich as a base for an additional treatment through which human communication and cognition can adequately be explicated.

It should be noted at this point that the concept of information has also been investigated in terms of the theory of computability, independent of the concept of uncertainty. In this approach, the amount of information represented by an object is measured by the length of the shortest possible program written in some standard language (e.g., a program for the standard Turing machine) by which the object is described in the sense that it can be computed. Information of this type is usually referred to as *descriptive information* to *algorithmic information*.

In this chapter, we are concerned solely with information conceived in terms of uncertainty reduction. To distinguish this conception of information from various other conceptions of information, let us call it *uncertainty-based information*.

The nature of uncertainty-based information depends on the mathematical theory within which uncertainty pertaining to various problem-solving situations is formalized. Each formalization of uncertainty in a problem-solving situation is a mathematical model of the situation. When we commit ourselves to a particular mathematical theory, our modelling becomes necessarily limited by the constraints of the theory. Clearly, a more general theory is capable of capturing uncertainties of some problem situations more faithfully than its less general competitors. As a rule, however, it involves greater computational demands.

Uncertainty-based information was first conceived in terms of classical set theory and, later, in terms of probability theory. The term *information theory* has almost invariably been used to refer to a theory based upon the well-known measure of probabilistic uncertainty established by Claude Shannon [1948]. Research on a broader conception of uncertainty-based information, liberated from the confines of classical set theory and probability theory, began in the early 1980s. The name *generalized information theory* was coined for a theory based upon this broader conception.

The ultimate goal of generalized information theory is to capture properties of uncertainty-based information formalized within any feasible mathematical framework. Although this goal has not been fully achieved as yet, substantial progress has been made in this direction since the early 1980s. In addition to classical set theory and probability theory, uncertainty-based information is now well understood in fuzzy set theory, possibility theory, and evidence theory.

When the seemingly unique connection between uncertainty and probability theory was broken, and uncertainty began to be conceived in terms of the much broader frameworks of fuzzy set theory and fuzzy measure theory, it soon became clear that uncertainty can be manifested in different forms. These forms represent distinct types of uncertainty. In probability theory, uncertainty is manifested only in one form.

Three types of uncertainty are now recognized in the five theories, in which measurement of uncertainty is currently well established. These three uncertainty types are: *nonspecificity* (or imprecision), which is connected with sizes (cardinalities) of relevant sets of alternatives; *fuzziness* (or vagueness), which results from imprecise boundaries of fuzzy sets; and *strife* (or discord), which expresses conflicts among the various sets of alternatives.

It is conceivable that other types of uncertainty will be discovered when the investigation of uncertainty extends to additional theories of uncertainty. Rather than speculating about this issue, this chapter is restricted to the three currently recognized types of uncertainty (and the associated information). It is shown, for each of the five theories of uncertainty, which uncertainty type is manifested in it and how the amount of uncertainty of that type can adequately be measured. The chapter is intended as a summary of existing results rather than a detailed exposition of the broad subject of uncertainty-based information. Hence, we do not cover axiomatic characterization of the various measures of uncertainty, proofs of their uniqueness, and other theoretical issues associated with them.

9.1 NONSPECIFICITY OF CRISP SETS

…urement of uncertainty (and associated information) was first conceived in terms of …ical set theory. It was shown by Hartley [1928] that using a function from the class of …ctions

$$U(A) = c \cdot \log_b |A|,$$

…ere $|A|$ denotes the cardinality of a finite nonempty set A, and b, c are positive constants …> 1, $c > 0$), is the only sensible way to measure the amount of uncertainty associated with …ite set of possible alternatives. Each choice of values of the constants b and c determines …e unit in which uncertainty is measured. When $b = 2$ and $c = 1$, which is the most common …oice, uncertainty is measured in *bits*, and we obtain

$$U(A) = \log_2 |A|. \tag{9.1}$$

…e bit of uncertainty is equivalent to the total uncertainty regarding the truth or falsity of …e proposition.

Let the set function U defined by (9.1) be called a *Hartley function*. Its uniqueness …s a measure of uncertainty (in bits) associated with sets of alternatives can also be proven …iomatically.

When the Hartley function U is applied to nonempty subsets of a given finite universal …et X, it has the form

$$U : \mathcal{P}(X) - \{\emptyset\} \to \mathbb{R}^+.$$

In this case, its range is

$$0 \leq U(A) \leq \log_2 |X|.$$

The meaning of uncertainty measured by the Hartley function depends on the meaning of the set A. For example, when A is a set of predicted states of a variable (from the set X of all states defined for the variable), $U(A)$ is a measure of *predictive uncertainty*; when A is a set of possible diseases of a patient determined from relevant medical evidence, $U(A)$ is a measure of *diagnostic uncertainty*; when A is a set of possible answers to an unsettled historical question, $U(A)$ is a measure of *retrodictive uncertainty*; when A is a set of possible policies, $U(A)$ is a measure of *prescriptive uncertainty*.

Observe that uncertainty expressed in terms of sets of alternatives results from the nonspecificity inherent in each set. Large sets result in less specific predictions, retrodictions, and so forth than their smaller counterparts. Full specificity is obtained when all alternatives are eliminated except one. Hence, uncertainty expressed by sets of possible alternatives and measured by the Hartley function is well characterized by the term *nonspecificity*.

Consider now a situation characterized by a set A of possible alternatives (predictive, prescriptive, etc.). Assume that this set is reduced to its subset B by some action. Then, the amount of uncertainty-based information, $I(A, B)$, produced by the action, which is relevant to the situation, is equal to the amount of reduced uncertainty given by the difference $U(A) - U(B)$. That is,

$$I(A, B) = \log_2 \frac{|A|}{|B|}. \tag{9.2}$$

When the action eliminates all alternatives except one (i.e., when $|B| = 1$), we have $I(A, B) = \log_2 |A| = U(A)$. This means that $U(A)$ may also be viewed as the amount of information needed to characterize one element of set A.

Consider now two universal sets, X and Y, and assume that a relation $R \subseteq X \times Y$ describes a set of possible joint alternatives in some situation of interest. Assume further that the domain and range of R are sets $R_X \subseteq X$ and $R_Y \subseteq Y$, respectively. Then three distinct Hartley functions are applicable, defined on the power sets of X, Y, and $X \times Y$. To identify clearly which universal set is involved in each case, it is useful (and a common practice) to write $U(X)$, $U(Y)$, $U(X, Y)$ instead of $U(R_X)$, $U(R_Y)$, and $U(R)$, respectively. Functions

$$U(X) = \log_2 |R_X|,$$

$$U(Y) = \log_2 |R_Y|$$

are called *simple uncertainties*, while function

$$U(X, Y) = \log_2 |R|$$

is called a *joint uncertainty*.

Two additional Hartley functions are defined by the formulas

$$U(X|Y) = \log_2 \frac{|R|}{|R_Y|},$$

$$U(Y|X) = \log_2 \frac{|R|}{|R_X|},$$

which are called *conditional uncertainties*.

Observe that the ratio $|R|/|R_Y|$ in $U(X|Y)$ represents the average number of elements of X that are possible alternatives under the condition that an element of Y has already been selected. This means that $U(X|Y)$ measures the average nonspecificity regarding alternative choices from X for all particular choices from Y. Function $U(Y|X)$ clearly has a similar meaning, with the roles of sets X and Y exchanged. Observe also that the conditional uncertainties can be expressed in terms of the joint uncertainty and the two simple uncertainties:

$$U(X|Y) = U(X, Y) - U(Y), \tag{9.8}$$

$$U(Y|X) = U(X, Y) - U(X). \tag{9.9}$$

Furthermore,

$$U(X) - U(Y) = U(X|Y) - U(Y|X), \tag{9.10}$$

which follows immediately from (9.8) and (9.9).

If possible alternatives from X do not depend on selections from Y, and vice versa, then $R = X \times Y$ and the sets X and Y are called *noninteractive*. Then, clearly,

$$U(X|Y) = U(X), \tag{9.11}$$

$$U(Y|X) = U(Y), \tag{9.12}$$

$$U(X, Y) = U(X) + U(Y). \tag{9.13}$$

In all other cases, when sets X and Y are *interactive*, these equations become the inequalities

$$U(X|Y) \leq U(X), \tag{9.14}$$

$$U(Y|X) \leq U(Y), \tag{9.15}$$

$$U(X,Y) \leq U(X) + U(Y). \tag{9.16}$$

The following symmetric function, which is usually referred to as *information transmission*, is a useful indicator of the strength of constraint between sets X and Y:

$$T(X,Y) = U(X) + U(Y) - U(X,Y). \tag{9.17}$$

When the sets are noninteractive, $T(X,Y) = 0$; otherwise, $T(X,Y) > 0$. Using (9.8) and (9.9), $T(X,Y)$ can also be expressed in terms of the conditional uncertainties

$$T(X,Y) = U(X) - U(X|Y), \tag{9.18}$$

$$T(X,Y) = U(Y) - U(Y|X). \tag{9.19}$$

Information transmission can be generalized to express the constraint among more than two sets. It is always expressed as the difference between the total information based on the individual sets and the joint information. Formally,

$$T(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} U(X_i) - U(X_1, X_2, \ldots, X_n). \tag{9.20}$$

Example 9.1

Consider two variables x and y whose values are taken from sets $X = \{\text{low, medium, high}\}$ and $Y = \{1, 2, 3, 4\}$, respectively. It is known that the variables are constrained by the relation R, expressed by the matrix

$$\begin{array}{l} \\ \text{Low} \\ \text{Medium} \\ \text{High} \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

We can see that the *low* value of x does not constrain y at all, the *medium* value of x constrains y partially, and the *high* value constrains it totally. The following types of Hartley information can be calculated in this example:

$$U(X) = \log_2 |X| = \log_2 3 = 1.6,$$

$$U(Y) = \log_2 |Y| = \log_2 4 = 2,$$

$$U(X,Y) = \log_2 |R| = \log_2 7 = 2.8,$$

$$U(X|Y) = U(X,Y) - U(Y) = 2.8 - 2 = .8,$$

$$U(Y|X) = U(X,Y) - U(X) = 2.8 - 1.6 = 1.2,$$

$$T(X,Y) = U(X) + U(Y) - U(X,Y) = 1.6 + 2 - 2.8 = .8.$$

The Hartley function in the form (9.1) is applicable only to finite sets. However, this form may be appropriately modified to infinite sets on $\mathbb{R}$ (or, more generally, $\mathbb{R}^k$ for some natural number k). Given a measurable and Lebesgue-integrable subset A of $\mathbb{R}$ (or $\mathbb{R}^k$), a meaningful counterpart of (9.1) for infinite sets takes the form

$$U(A) = \log[1 + \mu(A)],$$

where $\mu(A)$ is the measure of A defined by the Lebesgue integral of the characteristic function of A. For example, when A is an interval $[a, b]$ on $\mathbb{R}$, then $\mu(A) = b - a$ and

$$U([a, b]) = \log[1 + b - a].$$

The choice of the logarithm in (9.21) is less significant than in (9.1) since values $U(A)$ obtained for infinite sets by (9.21) do not yield any meaningful interpretation in terms of uncertainty regarding truth values of a finite number of propositions and, consequently, they are not directly comparable with values obtained for finite sets by (9.1). For its mathematical convenience, the natural logarithm is a suitable choice.

9.3 NONSPECIFICITY OF FUZZY SETS

A natural generalization of the Hartley function from classical set theory to fuzzy set theory was proposed in the early 1980s under the name *U-uncertainty*. For any nonempty fuzzy set A defined on a finite universal set X, the generalized Hartley function has the form

$$U(A) = \frac{1}{h(A)} \int_0^{h(A)} \log_2 |{}^{\alpha}A| \, d\alpha, \tag{9.22}$$

where $|{}^{\alpha}A|$ denotes the cardinality of the α-cut of A and $h(A)$ is the height of A. Observe that $U(A)$, which measures nonspecificity of A, is a weighted average of values of the Hartley function for all distinct α-cuts of the normalized counterpart of A, defined by $A(x)/h(A)$ for all $x \in X$. Each weight is a difference between the values of α of a given α-cut and the immediately preceding α-cut. For any $A, B \in \mathcal{F}(X) - \{\varnothing\}$, if $A(x)/h(A) = B(x)/h(B)$ for all $x \in X$, then $U(A) = U(B)$. That is, fuzzy sets that are equal when normalized have the same nonspecificity measured by function U.

Example 9.2

Consider a fuzzy set A on $\mathbb{N}$ whose membership function is defined by the dots in the diagram in Fig. 9.1; we assume that $A(x) = 0$ for all $x > 15$. Applying formula (9.22) to A, we obtain

$$\begin{aligned}
\int_0^1 \log_2 |{}^{\alpha}A| \, d\alpha &= \int_0^{.1} \log_2 15 \, d\alpha + \int_{.1}^{.3} \log_2 12 \, d\alpha + \int_{.3}^{.4} \log_2 11 \, d\alpha \\
&\quad + \int_{.4}^{.6} \log_2 9 \, d\alpha + \int_{.6}^{.7} \log_2 7 \, d\alpha + \int_{.7}^{.9} \log_2 5 \, d\alpha + \int_{.9}^{1} \log_2 3 \, d\alpha \\
&= .1 \log_2 15 + .2 \log_2 12 + .1 \log_2 11 + .2 \log_2 9 + .1 \log_2 7 \\
&\quad + .2 \log_2 5 + .1 \log_2 3 = 2.99.
\end{aligned}$$

The amount of nonspecifity associated with the given fuzzy set is thus approximately three bits. This is equivalent to the nonspecifity of a crisp set that contains eight elements, for example, the set $\{6, 7, \ldots, 13\}$, whose characteristic function is illustrated in Fig. 9.1 by the shaded area.

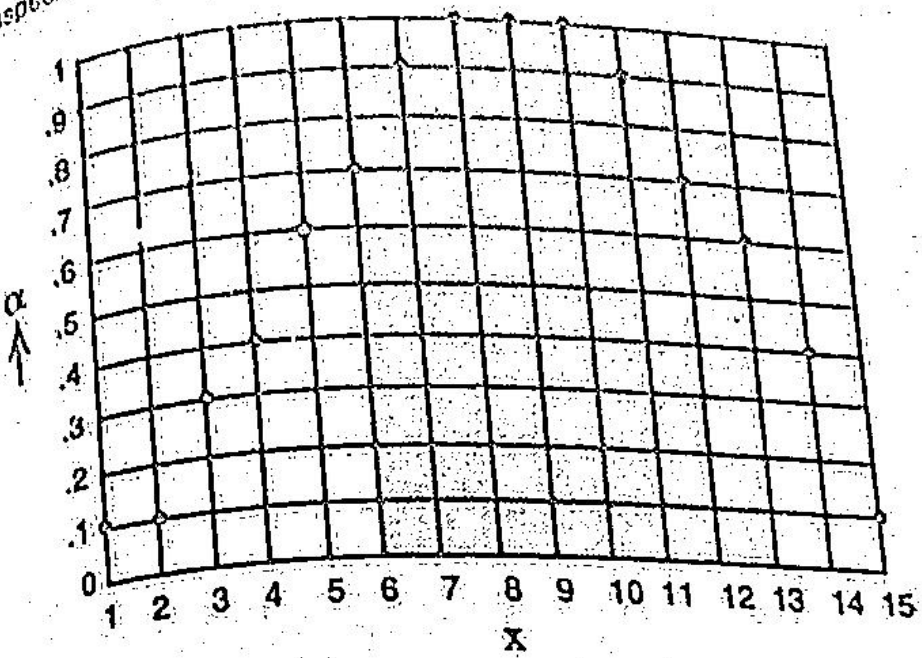


Figure 9.1 Membership functions of fuzzy set A (defined by the dots) and crisp set C (defined by the shaded area).

When a nonempty fuzzy set A is defined on $\mathbb{R}$ (or $\mathbb{R}^k$ for some $k \in \mathbb{N}$), and the α-cuts ${}^{\alpha}A$ infinite sets (e.g., intervals of real numbers), we have to calculate $U(A)$ by the modified m

$$U(A) = \frac{1}{h(A)} \int_0^{h(A)} \log[1 + \mu({}^{\alpha}A)]d\alpha, \tag{9.23}$$

hich is a generalization of (9.21). It is assumed that ${}^{\alpha}A$ is a measurable and Lebesgue-tegrable function; $\mu({}^{\alpha}A)$ is the measure of ${}^{\alpha}A$ defined by the Lebesgue integral of the haracteristic function of ${}^{\alpha}A$. As in the discrete case, defined by (9.22), fuzzy sets that are qual when normalized have the same nonspecificity.

Example 9.3

Consider fuzzy sets A_1, A_2, A_3 on $\mathbb{R}$, whose membership functions are depicted in Fig. 9.2. To calculate the nonspecificities of these fuzzy sets by (9.23), we have to determine for each set A_i $(i = 1, 2, 3)$ the measure $\mu({}^{\alpha}A_i)$ as a function of α. In each case, the α-cuts ${}^{\alpha}A_i$ are intervals ${}^{\alpha}A_i = [a_i(\alpha), b_i(\alpha)]$, and hence, $\mu({}^{\alpha}A_i) = b_i(\alpha) - a_i(\alpha)$. For A_1,

$$A_1(x) = \begin{cases} x/2 & \text{for } x \in [0, 2] \\ 2 - x/2 & \text{for } x \in [2, 4] \\ 0 & \text{otherwise,} \end{cases}$$

and $a_1(\alpha)$ and $b_1(\alpha)$ are determined by the equations $\alpha = a_1(\alpha)/2$ and $\alpha = 2 - b_1(\alpha)/2$. Hence, ${}^{\alpha}A_1 = [2\alpha, 4 - 2\alpha]$ and $\mu({}^{\alpha}A_1) = 4 - 4\alpha$. Now applying (9.23) and choosing the natural logarithm for convenience, we obtain

$$U(A_1) = \int_0^1 \ln(5 - 4\alpha)d\alpha = \left[-\frac{1}{4}(5 - 4\alpha)\ln(5 - 4\alpha) - \alpha\right]_0^1$$

$$= \frac{5}{4}\ln 5 - 1 = 1.012.$$

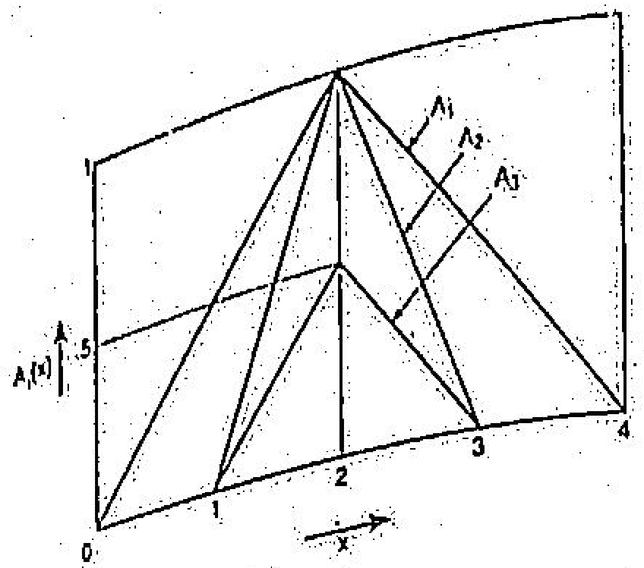


Figure 9.2 Fuzzy sets employed in Example 9.3.

Similarly, for the other two fuzzy sets, we have

$$U(A_2) = \int_0^1 \ln(3-2\alpha)d\alpha = \left[-\frac{1}{2}(3-2\alpha)\ln(3-2\alpha)-\alpha\right]_0^1$$

$$= \frac{3}{2}\ln 3 - 1 = 0.648,$$

$$U(A_3) = 2\int_0^{.5} \ln(3-4\alpha)d\alpha = 2\left[-\frac{1}{4}(3-4\alpha)\ln(3-4\alpha)-\alpha\right]_0^{.5}$$

$$= 2(\frac{3}{4}\ln 3 - \frac{1}{2}) = 0.648.$$

Observe that set A_3, which is not normal, can be normalized by replacing $A_3(x)$ with $2A_3(x)$ for each $x \in \mathbb{R}$. Since $2A_3(x) = A_2(x)$, set A_2 may be viewed as the normalized counterpart of A_3. These two sets have the same nonspecificity, which is true for any fuzzy set and its normal counterparts (the latter may, of course, be the set itself).

In general, given a normal fuzzy set $A \in \mathcal{F}(X)$, all fuzzy sets B in the infinite family

$$\mathcal{B}_A = \{B \in \mathcal{F}(X) - \{\varnothing\} \mid B(x) = h(B)A(x) \text{ for all } x \in X\} \tag{9.24}$$

have the same nonspecificity, regardless whether X is finite or infinite and we apply (9.22) or (9.23), respectively. Although the sets in $\mathcal{B}_A$ are not distinguished by the nonspecificity, they are clearly distinguished by their heights $h(B)$. For each fuzzy set $B \in \mathcal{B}_A$, its height may be viewed as the degree of *validity* or *credibility* of information expressed by the fuzzy set. From this point of view, information expressed by the empty set has no validity. This explains why it is reasonable to exclude the empty set from the domain of function U.

The U-uncertainty was investigated more thoroughly within possibility theory, utilizing ordered possibility distributions. In this domain, function U has the form

$$U : \mathcal{R} \to \mathbb{R}^+,$$

where $\mathcal{R}$ denotes the set of all finite and ordered possibility distributions, each of which represents a normal fuzzy set. Given a possibility distribution

$$\mathbf{r} = \langle r_1, r_2, \ldots, r_n \rangle \text{ such that } 1 = r_1 \geq r_2 \geq \ldots \geq r_n,$$

of the U-uncertainty of $\mathbf{r}$, $U(\mathbf{r})$, can be expressed by a convenient form,

$$U(\mathbf{r}) = \sum_{i=2}^{n} (r_i - r_{i+1}) \log_2 i, \tag{9.25}$$

where $r_{n+1} = 0$ by convention.

Assume that the possibility distribution $\mathbf{r}$ in (9.25) represents a normal fuzzy set A in (9.22) in the way discussed in Sec. 7.4. Then it is easy to see that $U(\mathbf{r}) = U(A)$ whenever $r_i - r_{i+1} > 0$, i represents the cardinality of the α-cut with $\alpha = r_i$.

When terms on the right-hand side of (9.25) are appropriately combined, the U-uncertainty can be expressed by another formula:

$$U(\mathbf{r}) = \sum_{i=2}^{n} r_i \log_2 \frac{i}{i-1}. \tag{9.26}$$

Furthermore, (9.25) can be written as

$$U(\mathbf{m}) = \sum_{i=2}^{n} m_i \log_2 i, \tag{9.27}$$

where $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ represents the basic probability assignment corresponding to $\mathbf{r}$ in the sense of the convention introduced in Sec. 7.3.

Example 9.4

Calculate $U(\mathbf{r})$ for the possibility distribution

$$\mathbf{r} = (1, 1, .8, .7, .7, .7, .4, .3, .2, .2).$$

Let us use (9.27), which is particularly convenient for calculating the U-uncertainty. First, using the equation $m_i = r_i - r_{i+1}$ for all $i = 1, 2, \ldots, 10$, and assuming $r_{11} = 0$, we determine

$$\mathbf{m} = \iota(\mathbf{r}) = (0, .2, .1, 0, 0, .3, .1, .1, 0, .2).$$

Now, we apply components of $\mathbf{m}$ to (9.27) and obtain

$$\begin{aligned} U(\mathbf{r}) &= 0 \log_2 1 + .2 \log_2 2 + .1 \log_2 3 + 0 \log_2 4 + 0 \log_2 5 + .3 \log_2 6 \\ &\quad + .1 \log_2 7 + .1 \log_2 8 + 0 \log_2 9 + .2 \log_2 10 = 2.18 \end{aligned}$$

Consider now two universal sets, X and Y, and a joint possibility distribution $\mathbf{r}$ defined on $X \times Y$. Adopting the notation introduced for the Hartley function, let $U(X, Y)$ denote the joint U-uncertainty, and let $U(X)$ and $U(Y)$ denote simple U-uncertainties defined on the marginal possibility distributions $\mathbf{r}_X$ and $\mathbf{r}_Y$, respectively. Then, we have

$$U(X) = \sum_{A \in \mathcal{F}_X} m_X(A) \log_2 |A|, \tag{9.28}$$

$$U(Y) = \sum_{B \in \mathcal{F}_Y} m_Y(B) \log_2 |B|, \tag{9.29}$$

$$U(X, Y) = \sum_{A \times B \in \mathcal{F}} m(A \times B) \log_2 |A \times B|, \tag{9.30}$$

where $\mathcal{F}_X, \mathcal{F}_Y, \mathcal{F}$ are sets of focal elements induced by m_X, m_Y, m, respectively. Furthermore, we define conditional U-uncertainties, $U(X|Y)$ and $U(Y|X)$, as the following generalizations of the corresponding conditional Hartley functions:

$$U(X|Y) = \sum_{A \times B \in \mathcal{F}} m(A \times B) \log_2 \frac{|A \times B|}{|B|}, \tag{9.31}$$

$$U(Y|X) = \sum_{A \times B \in \mathcal{F}} m(A \times B) \log_2 \frac{|A \times B|}{|A|}. \tag{9.32}$$

Observe that the term $|A \times B|/|B|$ in (9.31) represents for each focal element $A \times B$ in $\mathcal{F}$ the average number of elements of A that remain possible alternatives under the condition that an element of Y has already been selected. Expressing $U(X|Y)$ in the form of (9.22), we have

$$\begin{aligned} U(X|Y) &= \int_0^1 \log_2 \frac{|^\alpha(E \times F)|}{|^\alpha F|} d\alpha \\ &= \int_0^1 \log_2 |^\alpha(E \times F)| d\alpha - \int_0^1 \log_2 |^\alpha F| d\alpha \\ &= U(X, Y) - U(Y). \end{aligned} \tag{9.33}$$

This equation is clearly a generalization of (9.8) from crisp sets to normal fuzzy sets, and by (9.22) to subnormal fuzzy sets. Observe that the focal elements A, B in (9.31) correspond to the α-cuts $^\alpha E$, $^\alpha F$ in (9.33), respectively. In a similar way, a generalization of (9.9) can be defined. Using these two generalized equations, it can easily be shown that (9.10)–(9.16) are also valid for the U-uncertainty. Furthermore, information transmission can be defined for the U-uncertainty by (9.17), and, then, (9.18) and (9.19) are also valid in the generalized framework.

9.4 FUZZINESS OF FUZZY SETS

The second type of uncertainty that involves fuzzy sets (but not crisp sets) is fuzziness (or vagueness). In general, a *measure of fuzziness* is a function

$$f : \mathcal{F}(X) \to \mathbb{R}^+,$$

where $\mathcal{F}(X)$ denotes the set of all fuzzy subsets of X (fuzzy power set). For each fuzzy set A, this function assigns a nonnegative real number $f(A)$ that expresses the degree to which the boundary of A is not sharp.

In order to qualify as a sensible measure of fuzziness, function f must satisfy some requirements that adequately capture our intuitive comprehension of the degree of fuzziness. The following three requirements are essential:

1. $f(A) = 0$ iff A is a crisp set;
2. $f(A)$ attains its maximum iff $A(x) = 0.5$ for all $x \in X$, which is intuitively conceived as the highest fuzziness;
3. $f(A) \leq f(B)$ when set A is undoubtedly sharper than set B, which, according to our intuition, means that

$$A(x) \leq B(x) \text{ when } B(x) \leq 0.5$$

and

$$A(x) \geq B(x) \text{ when } B(x) \geq 0.5$$

for all $x \in X$.

There are different ways of measuring fuzziness that all satisfy the three essential requirements. One way is to measure fuzziness of any set A by a metric distance between its membership grade function and the membership grade function (or characteristic function) of the nearest crisp set. Even when committing to this conception of measuring fuzziness, the measurement is not unique. To make it unique, we have to choose a suitable distance function.

Another way of measuring fuzziness, which seems more practical as well as more general, is to view the fuzziness of a set in terms of the lack of distinction between the set and its complement. Indeed, it is precisely the lack of distinction between sets and their complements that distinguishes fuzzy sets from crisp sets. The less a set differs from its complement, the fuzzier it is. Let us restrict our discussion to this view of fuzziness, which is currently predominant in the literature.

Measuring fuzziness in terms of distinctions between sets and their complements is dependent on the definition of a fuzzy complement (Sec. 3.2). This issue is not discussed here. For the sake of simplicity, we assume that only the standard fuzzy complement is employed. If other types of fuzzy complements were also considered, the second and third of the three properties required of function f would have to be generalized by replacing the value 0.5 with the equilibrium of the fuzzy complement employed.

Employing the standard fuzzy complement, we can still choose different distance functions to express the lack of distinction of a set and its complement. One that is simple and intuitively easy to comprehend is the Hamming distance, defined by the sum of absolute values of differences. Choosing the Hamming distance, the local distinction (one for each $x \in X$) of a given set A and its complement is measured by

$$|A(x) - (1 - A(x))| = |2A(x) - 1|,$$

and the lack of each local distinction is measured by

$$1 - |2A(x) - 1|.$$

The measure of fuzziness, $f(A)$, is then obtained by adding all these local measurements:

$$f(A) = \sum_{x \in X} (1 - |2A(x) - 1|). \tag{9.34}$$

The range of function f is $[0, |X|]$; $f(A) = 0$ iff A is a crisp set; $f(A) = |X|$ when $A(x) = 0.5$ for all $x \in X$.

Applying, for example, (9.34) to the fuzzy set A defined in Fig. 9.1, we obtain

$$\begin{aligned} f(A) &= \sum_{i=1}^{15} (1 - |2A(x) - 1|) \\ &= 15 - (.8 + .8 + .4 + .2 + .2 + .4 + .8 + 1 + 1 + 1 + .8 + .4 + .2 + .2 + .8) \\ &= 15 - 9 = 6. \end{aligned}$$

Formula (9.34) is applicable only to fuzzy sets defined on finite universal sets. However, it can be readily modified to fuzzy sets defined on infinite but bounded subsets of $\mathbb{R}^n$ (for some natural number n). Consider, for example, that $X = [a, b]$, which is perhaps the most practical case. Then, (9.34) needs to be modified by replacing the summation with integration. This replacement results in the formula

$$f(A) = \int_a^b (1 - |2A(x) - 1|)dx$$

$$= b - a - \int_a^b |2A(x) - 1|dx. \qquad (9.35)$$

Example 9.5

Calculate the degrees of fuzziness of the three fuzzy sets defined in Fig. 9.2. Assuming that the sets are defined within the interval $[0, 4]$ and using (9.35), we obtain:

$$f(A_1) = 4 - \int_0^2 |x - 1|dx - \int_2^4 |3 - x|dx$$

$$= 4 - \int_0^1 (1 - x)dx - \int_1^2 (x - 1)dx - \int_2^3 (3 - x)dx - \int_3^4 (x - 3)dx$$

$$= 4 + \left[\frac{(1-x)^2}{2}\right]_0^1 - \left[\frac{(x-1)^2}{2}\right]_1^2 + \left[\frac{(3-x)^2}{2}\right]_2^3 - \left[\frac{(x-3)^2}{2}\right]_3^4$$

$$= 4 - .5 - .5 - .5 - .5 = 2,$$

$$f(A_2) = 4 - \int_0^1 dx - \int_1^2 |2x - 3|dx - \int_2^3 |5 - 2x|dx - \int_3^4 dx$$

$$= 4 - 1 - .5 - .5 - 1 = 1,$$

$$f(A_3) = 4 - \int_0^1 dx - \int_1^2 |x - 2|dx - \int_2^3 |2 - x|dx - \int_3^4 dx$$

$$= 4 - 1 - .5 - .5 - 1 = 1.$$

Although set A_3 and its normal counterpart, set A_2, happen to be equally fuzzy, this is not a general property, contrary to nonspecificity, of sets in a given family $\mathcal{B}_A$ defined by (9.21). To illustrate this difference between nonspecificity and fuzziness, let us calculate the degree of fuzziness for two additional fuzzy sets, A_4 and A_5, both of which belong to $\mathcal{B}_{A_2}$ with $h(A_4) = .4$ and $h(A_5) = .625$. That is,

$$A_k(x) = \begin{cases} h(A_k)(x - 1) & \text{for } x \in [1, 2] \\ h(A_k)(3 - x) & \text{for } x \in [2, 3] \\ 0 & \text{otherwise,} \end{cases}$$

where $k = 4, 5$, and we obtain:

$$f(A_4) = 4 - \int_0^1 dx - \int_1^2 (1.8 - .8x)dx - \int_2^3 (.8x - 1.4)dx - \int_3^4 dx$$

$$= 4 - 1 - .6 - .6 - 1 = .8,$$

$$f(A_5) = 4 - \int_0^1 dx - \int_1^2 |1.25x - 2.25|dx - \int_2^3 |2.75 - 1.25x|dx - \int_3^4 dx$$
$$= 4 - 1 - .425 - .425 - 1 = 1.15.$$

The meaning of the calculations in this example is illustrated in Fig. 9.3. For the universal set $[0, 4]$ and each of the fuzzy sets A_1–A_5, the shaded areas indicate for each $x \in [0, 4]$ the difference between membership grades of the set and its complement. Since the degree of fuzziness is expressed in terms of the deficiency of these differences with respect to 1, it is measured by the total size of the unshaded areas in each diagram.

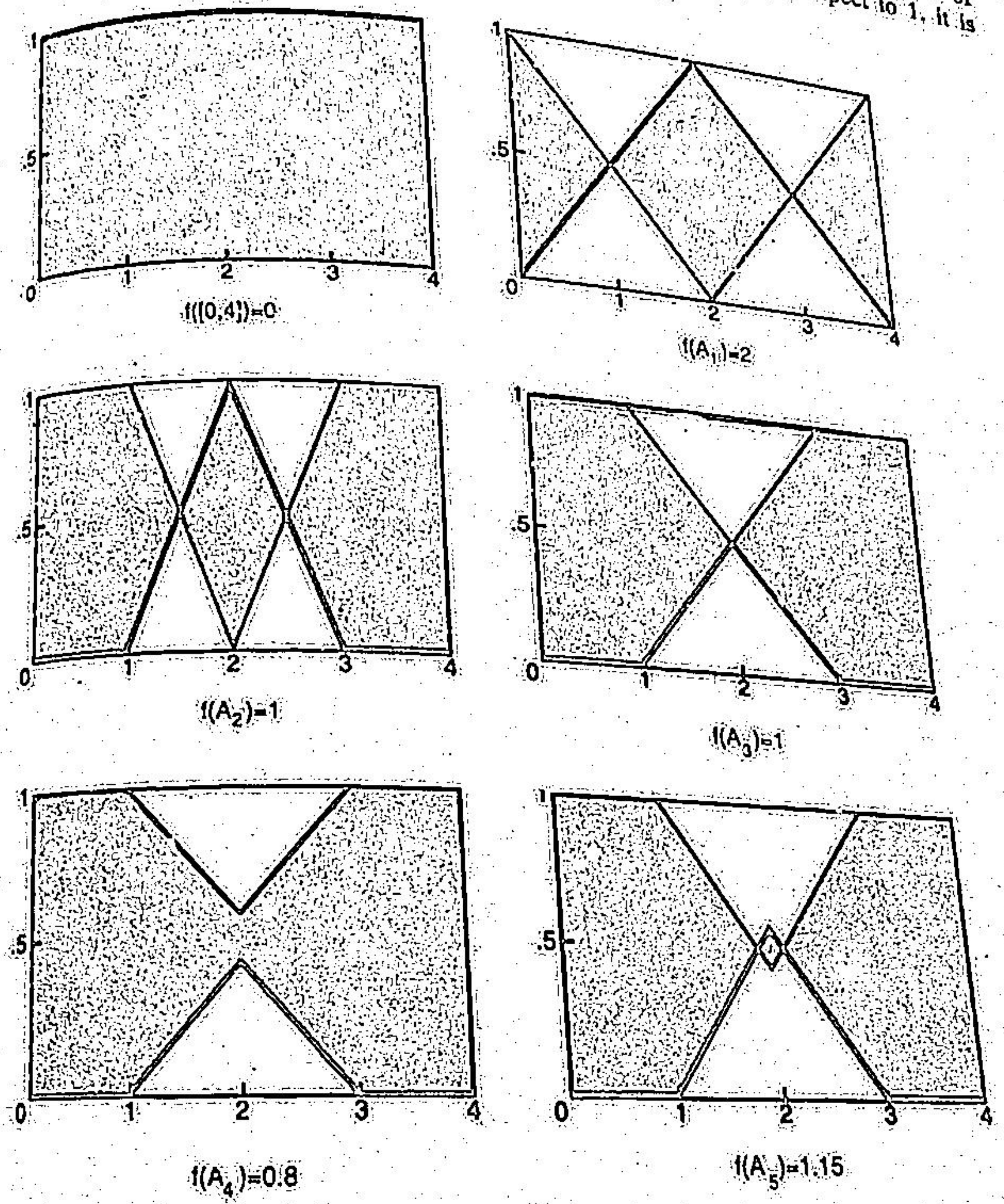


Figure 9.3 Illustration of the meaning of (9.35) in the context of Example 9.5.

It is important to realize that nonspecificity and fuzziness, which are both applica fuzzy sets, are distinct types of uncertainty. Moreover, they are totally independent of other. Observe, for example, that the two fuzzy sets defined by their membership func in Fig. 9.1 have almost the same nonspecificity (Example 9.2), but their degrees of fuzz are quite different. On the other hand, fuzzy sets depicted in Fig. 9.4 have very diffe nonspecificities, but their degrees of fuzziness are exactly the same.

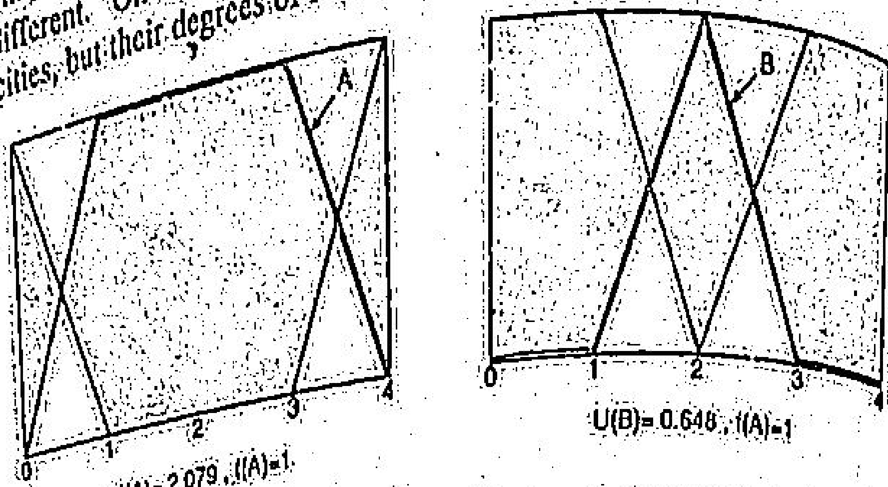


Figure 9.4 Example of a fuzzy set with very different nonspecificities, but equal fuzziness.

Fuzziness and nonspecificity are also totally different in their connections to information. When nonspecificity is reduced, we view the reduction as a gain in information, regardless of the associated change in fuzziness. The opposite, however, is not true. Whether it is meaningful to view a reduction in fuzziness as a gain in information depends on the accompanied change in nonspecificity. The view is reasonable when nonspecificity also decreases or remains the same, but it is questionable when nonspecificity increases. To illustrate this point, let us consider an integer-valued variable whose exact value is of interest, but the only available information regarding its value is given in terms of the fuzzy set A whose membership function is defined by the dots in Fig. 9.1. This information is incomplete; consequently, it results in uncertainty regarding the actual value of the variable: $U(A) = 2.99$, $f(A) = 6$. Assume now that we learn, as a result of some action, that the actual value of the variable is at least 6 and no more than 13. This defines a crisp set $B = \{6, 7, \ldots, 13\}$ whose characteristic function is illustrated in Fig. 9.1 by the shaded area. We have now $U(B) = 3$ and $f(B) = 0$. Although there is virtually no change in nonspecificity, the action helped us to completely eliminate fuzziness. It is reasonable to view the amount of eliminated fuzziness as a result of some proportional amount of information gained.

9.5 UNCERTAINTY IN EVIDENCE THEORY

Nonspecificity

The Hartley function, as a measure of nonspecificity, was first generalized from classical set theory to fuzzy set theory and possibility theory (Sec. 9.3). Once this generalized function, the U-uncertainty, was well established in possibility theory, a special branch of evidence theory, it was relatively easy to generalize it further and make it applicable within all of evidence theory.

To distinguish the U-uncertainty from its more general counterpart in evidence theory, the latter be denoted by N. It is defined by the formula

$$N(m) = \sum_{A \in \mathcal{F}} m(A) \log_2 |A|, \tag{9.36}$$

where $\langle \mathcal{F}, m \rangle$ is an arbitrary body of evidence. This function was proven a unique measure of nonspecificity in evidence theory under well-justified axioms and the choice of bits as measurement units. When focal elements in $\mathcal{F}$ are nested, N becomes the U-uncertainty.

Function N is clearly a weighted average of the Hartley function for all focal elements. The weights are values of the basic probability assignment. For each focal element A, $m(A)$ indicates the degree of evidence focusing on A, while $\log_2 |A|$ indicates the lack of specificity of this evidential claim. The larger the value of $m(A)$, the stronger the evidence; the larger the set A (and $\log_2 |A|$), the less specific the evidence. Consequently, it is reasonable to view function N as a *measure of nonspecificity*.

The range of function N is, as expected, $[0, \log_2 |X|]$. The minimum, $N(m) = 0$, is obtained when $m(\{x\}) = 1$ for some $x \in X$ (no uncertainty); the maximum, $N(m) = \log_2 |X|$, is reached for $m(X) = 1$ (total ignorance). It can easily be shown that (9.8)–(9.19) for the Hartley function remain valid when it is generalized to the N-uncertainty.

Since focal elements in probability measures are singletons, $|A| = 1$ and $\log_2 |A| = 0$ for each focal element. Consequently, $N(m) = 0$ for every probability measure. That is, probability theory is not capable of incorporating nonspecificity, one of the basic types of uncertainty. All probability measures are fully specific and, hence, are not distinguished from one another by their nonspecificities. What, then, is actually measured by the Shannon entropy, which is a well-established measure of uncertainty in probability theory? Before attempting to answer this question, let us review the key properties of the Shannon entropy.

The Shannon Entropy

The *Shannon entropy*, H, which is applicable only to probability measures, assumes in evidence theory the form

$$H(m) = -\sum_{x \in X} m(\{x\}) \log_2 m(\{x\}). \tag{9.37}$$

This function, which forms the basis of *classical information theory*, measures the average uncertainty (in bits) associated with the prediction of outcomes in a random experiment; its range is $[0, \log_2 |X|]$. Clearly, $H(m) = 0$ when $m(\{x\}) = 1$ for some $x \in X$; $H(m) = \log_2 |X|$ when m defines the uniform probability distribution on X (i.e., $m(\{x\}) = 1/|X|$ for all $x \in X$).

As the name suggests, function H was proposed by Shannon [1948]. It was proven in numerous ways, from several well-justified axiomatic characterizations, that this function is the only sensible measure of uncertainty in probability theory. It is also well known that (9.8)–(9.19) are valid when function U is replaced with function H.

Since values $m(\{x\})$ are required to add to 1 for all $x \in X$, (9.37) can be rewritten as

$$H(m) = -\sum_{x \in X} m(\{x\}) \log_2 [1 - \sum_{y \neq x} m(\{y\})]. \tag{9.38}$$

The term

$$\text{Con}(\{x\}) = \sum_{y \neq x} m(\{y\})$$

in (9.38) represents the total evidential claim pertaining to focal elements that are different with the focal element $\{x\}$. That is, Con $(\{x\})$ expresses the sum of all evidential claims that fully conflict with the one focusing on $\{x\}$. Clearly, Con $(\{x\}) \in [0,1]$ for each $x \in X$. The function $-\log_2[1 - \text{Con}(\{x\})]$, which is employed in (9.38), is monotonic increasing with Con $(\{x\})$ and extends its range from $[0,1]$ to $[0,\infty)$. The choice of the logarithmic function is a result of the axiomatic requirement that the joint uncertainty of several independent random variables be equal to the sum of their individual uncertainty.

It follows from these facts and from the form of (9.38) that the Shannon entropy is the mean (expected) value of the *conflict* among evidential claims within a given probabilistic body of evidence.

When a probability measure is defined on a real interval $[a,b]$ by a probability density function f, the Shannon entropy is not directly applicable as a measure of uncertainty. It can be employed only in a modified form,

$$D(f(x), g(x) | x \in [a,b]) = \int_a^b f(x) \log_2 \frac{f(x)}{g(x)} dx, \tag{9.40}$$

which involves two probability density functions, $f(x)$ and $g(x)$, defined on $[a,b]$. Its finite counterpart is the function

$$D(p(x), q(x) | x \in X) = \sum_{x \in X} p(x) \log_2 \frac{p(x)}{q(x)}, \tag{9.41}$$

which is known in information theory as the *Shannon cross-entropy* or *directed divergence*. Function D measures uncertainty in relative rather than absolute terms.

When $f(x)$ in (9.40) is replaced with a density function, $f(x,y)$, of a joint probability distribution on $X \times Y$, and $g(x)$ is replaced with the product of density functions of marginal distributions on X and Y, $f_X(x) \cdot f_Y(y)$, D becomes equivalent to the information transmission given by (9.17). This means that the continuous counterpart of the information transmission can be expressed as

$$D(f(x,y), f_X(x) \cdot f_Y(y) | x \in [a,b], y \in [c,d]) =$$

$$\int_c^d \int_a^b f(x,y) \log_2 \frac{f(x,y)}{f_X(x) \cdot f_Y(y)} dx dy. \tag{9.42}$$

Entropy-like Measure In Evid

...al Uncertainty

...ce the two types of uncertainty, nonspecificity and strife, coexist in evidence theory, and ...h are measured in the same units, it is reasonable to consider the possibility of adding their ...ividual measures to form a measure of total uncertainty. Following this idea and choosing ...e measure of strife defined by (9.54) as the best justified generalization of the Shannon ...ropy, the total uncertainty, NS, is defined by the equation

$$NS(m) = N(m) + S(m).$$

...stituting for $S(m)$ from (9.55) and for $N(m)$ from (9.36), we obtain

$$NS(m) = 2\sum_{A \in \mathcal{F}} m(A)\log_2|A| - \sum_{A \in \mathcal{F}} m(A)\log_2 \sum_{B \in \mathcal{F}} m(B)|A \cap B|$$

...r, in a more compact form,

$$NS(m) = \sum_{A \in \mathcal{F}} m(A)\log_2 \frac{|A|^2}{\sum_{B \in \mathcal{F}} m(B)|A \cap B|}. \tag{9.59}$$

...n possibility theory, the measure of total uncertainty clearly assumes the form

$$NS(\mathbf{r}) = \sum_{i=2}^{n} (r_i - r_{i+1})\log_2 \frac{i^2}{\sum_{j=1}^{i} r_j}. \tag{9.60}$$

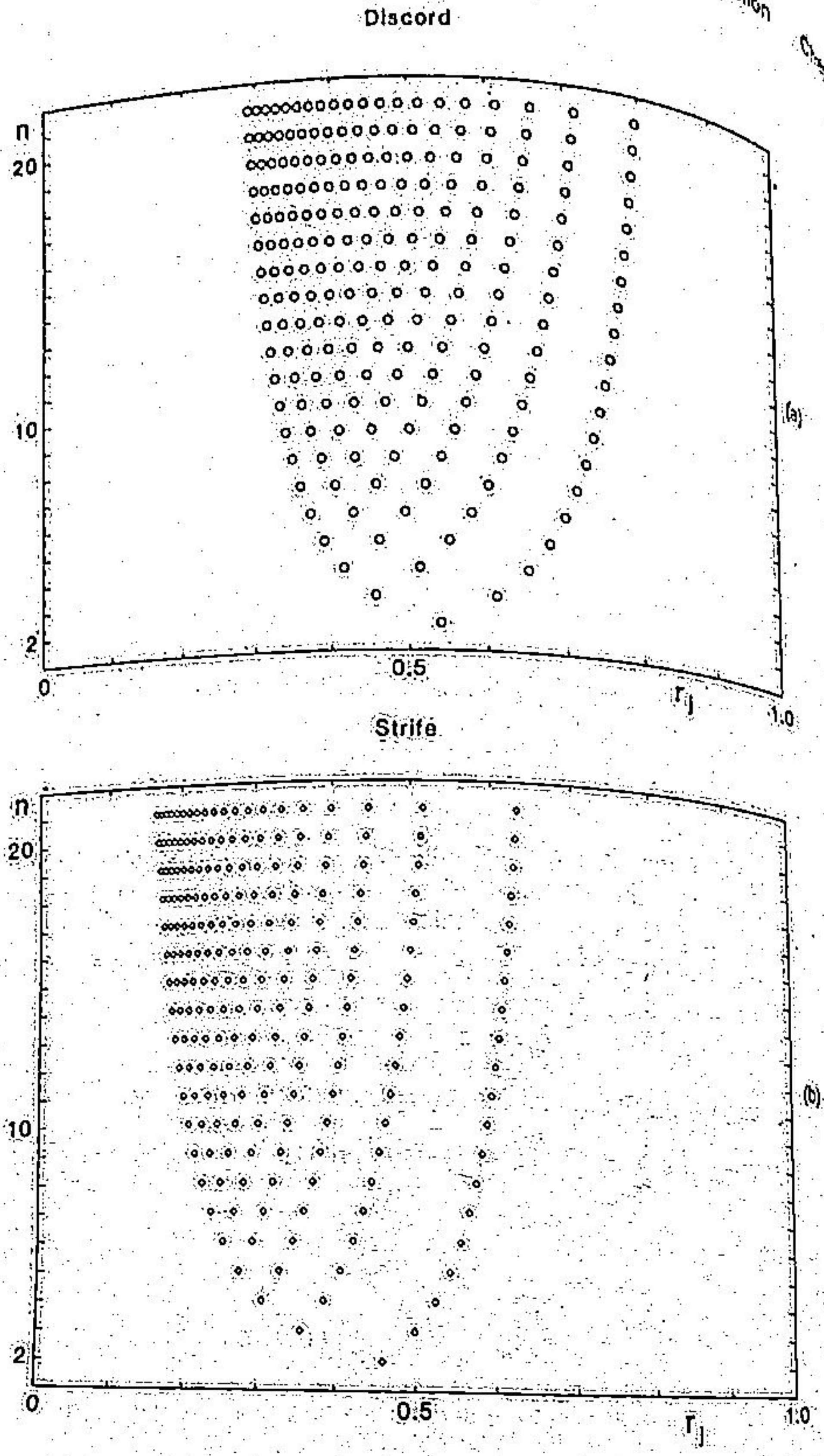
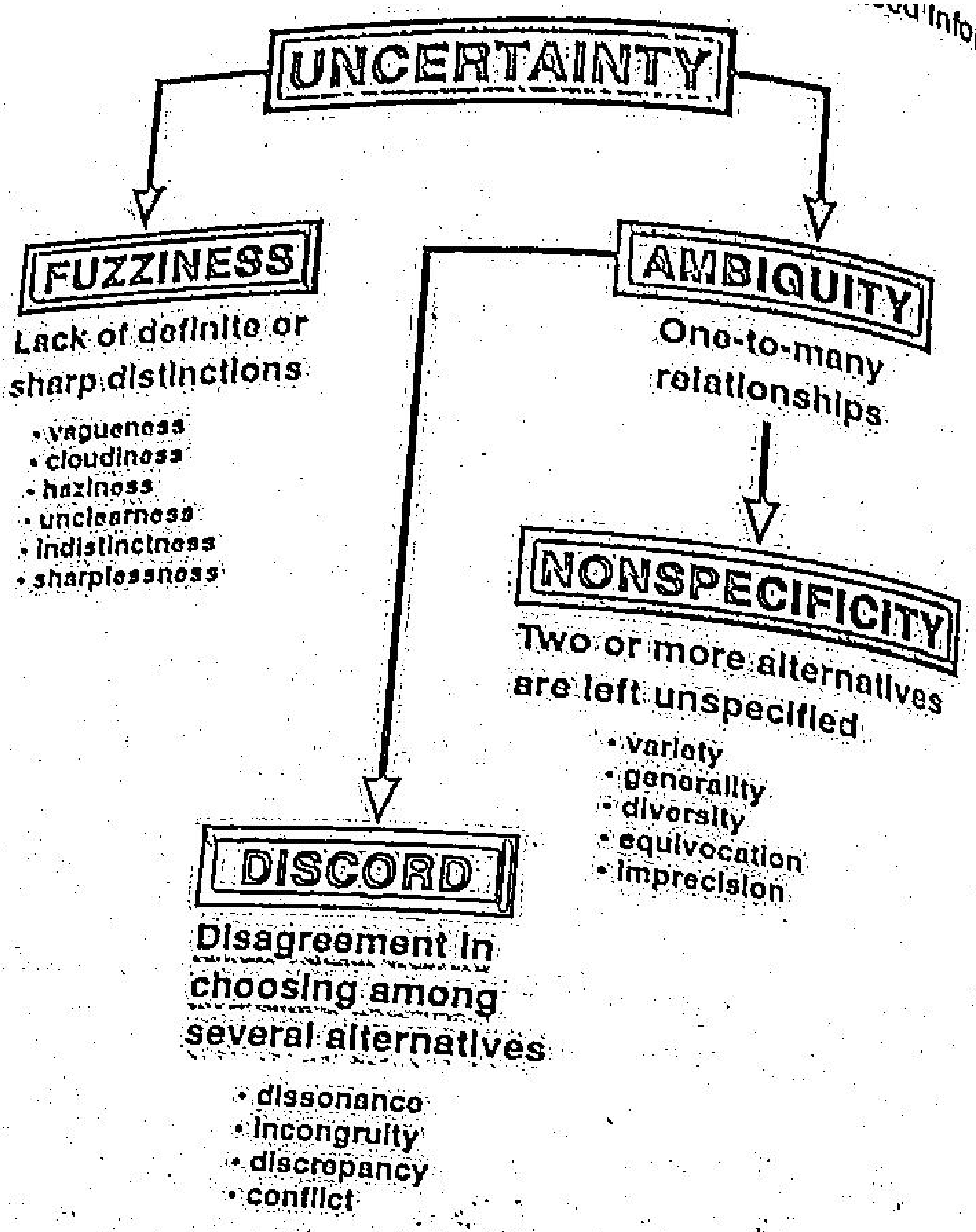


Figure 9.6 Values $r_1, r_2, \ldots, r_n$ for which the maxima shown in Fig. 9.5 are obtained: (a) discord; (b) strife.


UNCERTAINTY
FUZZINESS
Lack of definite or sharp distinctions
• vagueness
• cloudiness
• haziness
• unclearness
• indistinctness
• sharplessness
AMBIGUITY
One-to-many relationships
NONSPECIFICITY
Two or more alternatives are left unspecified
• variety
• generality
• diversity
• equivocation
• imprecision
DISCORD
Disagreement in choosing among several alternatives
• dissonance
• incongruity
• discrepancy
• conflict

PRINCIPLES OF UNCERTAINTY

...uncertainty (and information) measures become well justified, they can very effectively ...lized for managing uncertainty and the associated information. For example, they can ...lized for extrapolating evidence, assessing the strength of relationship between given ...s of variables, assessing the influence of given input variables on given output variables, ...suring the loss of information when a system is simplified, and the like. In many problem ...tions, the relevant measures of uncertainty are applicable only in their conditional or ...tive terms.

Although the utility of relevant uncertainty measures is as broad as the utility of any relevant measuring instrument, their role is particularly significant in three fundamental principles for managing uncertainty: the principle of minimum uncertainty, the principle of maximum uncertainty, and the principle of uncertainty invariance. Since types and measures of uncertainty substantially differ in different uncertainty theories, the principles result in considerably different mathematical problems when we move from one theory to another. In this section, we only explain the principal characteristics of each of these principles.

The Principle of Minimum Uncertainty

The *principle of minimum uncertainty* is basically an arbitration principle. It is used, in general, for narrowing down solutions in various systems problems that involve uncertainty. The principle states that we should accept only those solutions, from among all otherwise equivalent solutions, whose uncertainty (pertaining to the purpose concerned) is minimal.

A major class of problems for which the principle of minimum uncertainty is applicable are *simplification problems*. When a system is simplified, the loss of some information contained in the system is usually unavoidable. The amount of information that is lost in this process results in the increase of an equal amount of relevant uncertainty. Examples of relevant uncertainties are predictive, retrodictive, or prescriptive uncertainty. A sound simplification of a given system should minimize the loss of relevant information (or the increase in relevant uncertainty) while achieving the required reduction of complexity. That is, we should accept only such simplifications of a given system at any desirable level of complexity for which the loss of relevant information (or the increase in relevant uncertainty) is minimal. When properly applied, the principle of minimum uncertainty guarantees that no information is wasted in the process of simplification.

There are many simplification strategies, which can perhaps be classified into three main classes:

1. simplifications made by eliminating some entities from the system (variables subsystems, etc.);
2. simplifications made by aggregating some entities of the system (variables, states, etc.);
3. simplifications made by breaking overall systems into appropriate subsystems.

Regardless of the strategy employed, the principle of minimum uncertainty is utilized in the same way. It is an arbiter that decides which simplifications to choose at any given level of complexity.

Another application of the principle of minimum uncertainty is the area of *conflict-resolution problems*. For example, when we integrate several overlapping models into one larger model, the models may be locally inconsistent. It is reasonable, then, to require that each of the models be appropriately adjusted in such a way that the overall model becomes consistent. It is obvious that some information contained in the given models is inevitably lost by these adjustments. This is not desirable. Hence, we should minimize this loss of information. That is, we should accept only those adjustments for which the total loss of information (or total increase of uncertainty) is minimal. The total loss of information may be expressed, for example, by the sum of all individual losses or by a weighted sum, if the given models are valued differently.

9.1 Principle of Maximum Uncertainty

…cond principle, the *principle of maximum uncertainty*, is essential for any problem that …ves *ampliative reasoning*. This is reasoning in which conclusions are not entailed in …ven premises. Using common sense, the principle may be expressed by the following …ement: in any ampliative inference, use all information available, but make sure …no additional information is unwittingly added. That is, employing the connection …en information and uncertainty, the principle requires that conclusions resulting from …ampliative inference maximize the relevant uncertainty within the constraints representing …premises. This principle guarantees that our ignorance be fully recognized when we try …enlarge our claims beyond the given premises and, at the same time, that all information …ained in the premises be fully utilized. In other words, the principle guarantees that …conclusions are maximally noncommittal with regard to information not contained in the …mises.

Ampliative reasoning is indispensable to science in a variety of ways. For example, …enever we utilize a scientific model for predictions, we employ ampliative reasoning. …ilarly, when we want to estimate microstates from the knowledge of relevant macrostates and …ial information regarding the microstates (as in image processing and many other problems), …must resort to ampliative reasoning. The problem of the identification of an overall system …m some of its subsystems is another example that involves ampliative reasoning.

Ampliative reasoning is also common and important in our daily life, where, unfortu…ly, the principle of maximum uncertainty is not always adhered to. Its violation leads al…st invariably to conflicts in human communication, as well expressed by Bertrand Rus…l in his *Unpopular Essays* (London, 1950): "...whenever you find yourself getting an…y about a difference in opinion, be on your guard; you will probably find, on examina…n, that your belief is getting beyond what the evidence warrants."

The principle of maximum uncertainty is well developed and broadly utilized within …ssical information theory, where it is called the *principle of maximum entropy*. A general …mulation of the principle of maximum entropy is: determine a probability distribution …$(x)|x \in X$) that maximizes the Shannon entropy subject to given constraints $c_1, c_2, \ldots$, …ich express partial information about the unknown probability distribution, as well as general …straints (axioms) of probability theory. The most typical constraints employed in practical …plications of the maximum entropy principle are mean (expected) values of one or more …dom variables or various marginal probability distributions of an unknown joint distribution.

As an example, consider a random variable x with possible (given) nonnegative real …ues $x_1, x_2, \ldots, x_n$. Assume that probabilities $p(x_i)$ are not known, but we know the mean …pected) value $E(x)$ of the variable, which is related to the unknown probabilities by the …mula

$$E(x) = \sum_{i=1}^{n} x_i p(x_i). \tag{9.62}$$

…ploying the maximum entropy principle, we estimate the unknown probabilities $p(x_i)$, …$= 1, 2, \ldots, n$, by solving the following optimization problem: maximize function

$$H(p(x_i)|i = 1, 2, \ldots, n) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)$$

subject to the axiomatic constraints of probability theory and an additional constraint expressed by (9.62), where $E(x)$ and $x_1, x_2, \ldots, x_n$ are given numbers. When solving this problem, we obtain

$$p(x_i) = \frac{e^{-\beta x_i}}{\sum_{k=1}^{n} e^{-\beta x_k}}, \tag{9.63}$$

for all $i = 1, 2, \ldots, n$, where β is a constant obtained by solving (numerically) the equation

$$\sum_{i=1}^{n} [x_i - E(x)] e^{-\beta[x_i - E(x)]} = 0. \tag{9.64}$$

Our only knowledge about the random variable x in this example is the knowledge of its expected value $E(x)$. It is expressed by (9.62) as a constraint on the set of relevant probability distributions. If $E(x)$ were not known, we would be totally ignorant about x, and the maximum entropy principle would yield the uniform probability distribution (the only distribution for which the entropy reaches its absolute maximum). The entropy of the probability distribution given by (9.63) is smaller than the entropy of the uniform distribution, but it is the largest entropy from among all the entropies of the probability distributions that conform to the given expected value $E(x)$.

A generalization of the principle of maximum entropy is the *principle of minimum cross-entropy*. It can be formulated as follows: given a prior probability distribution function q on a finite set X and some relevant new evidence, determine a new probability distribution function p that minimizes the cross-entropy D given by (9.41) subject to constraints $c_1, c_2, \ldots$, which represent the new evidence, as well as to the standard constraints of probability theory.

New evidence reduces uncertainty. Hence, uncertainty expressed by p is, in general, smaller than uncertainty expressed by q. The principle of minimum cross-entropy helps us to determine how much smaller it should be. It allows us to reduce the uncertainty of q in the smallest amount necessary to satisfy the new evidence. That is, the posterior probability distribution function p estimated by the principle is the largest among all other distribution functions that conform to the evidence.

Optimization problems that emerge from the maximum uncertainty principle outside classical information theory have yet to be properly investigated and tested in praxis. When several types of uncertainty are applicable, we must choose one from several possible optimization problems. In evidence theory, for example, the principle of maximum uncertainty yields four possible optimization problems, which are distinguished from one another by the objective function involved: nonspecificity, strife, total uncertainty, or both nonspecificity and strife viewed as two distinct objective functions.

As a simple example to illustrate the principle of maximum nonspecificity in evidence theory, let us consider a finite universal set X, three nonempty subsets of which are of interest to us: A, B, and $A \cap B$. Assume that the only evidence on hand is expressed in terms of two numbers, a and b, that represent the total beliefs focusing on A and B, respectively ($a, b \in [0, 1]$). Our aim is to estimate the degree of support for $A \cap B$ based on this evidence.

As a possible interpretation of this problem, let X be a set of diseases considered in an expert system designed for medical diagnosis in a special area of medicine, and let A and B be sets of diseases that are supported for a particular patient by some diagnostic tests to

a and b, respectively. Using this evidence, it is reasonable to estimate the degree of support for diseases in $A \cap B$ by using the principle of maximum nonspecificity. This principle is a safeguard which does not allow us to produce an answer (diagnosis) that is more specific than warranted by the evidence.

The use of the principle of maximum nonspecificity leads, in our example, to the following optimization problem:

Determine values $m(X)$, $m(A)$, $m(B)$, and $m(A \cap B)$ for which the function $m(X)\log_2|X| + m(A)\log_2|A| + m(B)\log_2|B| + m(A \cap B)\log_2|A \cap B|$ reaches its maximum subject to the constraints

$$m(A) + m(A \cap B) = a,$$
$$m(B) + m(A \cap B) = b,$$
$$m(X) + m(A) + m(B) + m(A \cap B) = 1,$$
$$m(X), m(A), m(B), m(A \cap B) \geq 0,$$

where $a, b \in [0, 1]$ are given numbers.

The constraints are represented in this case by three linear algebraic equations of four unknowns and, in addition, by the requirement that the unknowns be nonnegative real numbers. The first two equations represent our evidence; the third equation and the inequalities represent general constraints of evidence theory. The equations are consistent and independent. Hence, they involve one degree of freedom. Selecting, for example, $m(A \cap B)$ as the free variable, we readily obtain

$$\begin{aligned} m(A) &= a - m(A \cap B), \\ m(B) &= b - m(A \cap B), \\ m(X) &= 1 - a - b + m(A \cap B). \end{aligned} \tag{9.65}$$

Since all the unknowns must be nonnegative, the first two equations set the upper bound of $m(A \cap B)$, whereas the third equation specifies its lower bound; the bounds are

$$\max(0, a + b - 1) \leq m(A \cap B) \leq \min(a, b). \tag{9.66}$$

Using (9.65), the objective function can now be expressed solely in terms of the free variable $m(A \cap B)$. After a simple rearrangement of terms, we obtain

$$m(A \cap B)[\log_2|X| - \log_2|A| - \log_2|B| + \log_2|A \cap B|]$$
$$+ (1 - a - b)\log_2|X| + a\log_2|A| + b\log_2|B|.$$

Clearly, only the first term in this expression can influence its value, so we may rewrite the expression as

$$m(A \cap B)\log_2 K_1 + K_2, \tag{9.67}$$

where

$$K_1 = \frac{|X| \cdot |A \cap B|}{|A| \cdot |B|}$$

and

$$K_2 = (1 - a - b)\log_2 |X| + a \log_2 |A| + b \log_2 |B|$$

are constant coefficients. The solution to the optimization problem depends only on the values of K_1. Since A, B, and $A \cap B$ are assumed to be nonempty subsets of X, $K_1 > 0$. If $K_1 < 1$, then $\log_2 K_1 < 0$, and we must minimize $m(A \cap B)$ to obtain the maximum of expression (9.67); hence, $m(A \cap B) = \max(0, a + b - 1)$ due to (9.66). If $K_1 > 1$, then $\log_2 K_1 > 0$, and we must maximize $m(A \cap B)$; hence, $m(A \cap B) = \min(a, b)$, as given by (9.66). When $K_1 = 1$, $\log_2 K_1 = 0$, and expression (9.67) is independent of $m(A \cap B)$; this implies that the solution is not unique or, more precisely, that any value of $m(A \cap B)$ in the range (9.66) is a solution to the optimization problem. The complete solution can thus be expressed by the following equations:

$$m(A \cap B) = \begin{cases} \max(0, a + b - 1) & \text{when } K_1 < 1 \\ [\max(0, a + b - 1), \min(a, b)] & \text{when } K_1 = 1 \\ \min(a, b) & \text{when } K_1 > 1. \end{cases} \quad (9.68)$$

The three types of solutions are exemplified in Table 9.2.

TABLE 9.2 EXAMPLES OF THE THREE TYPES OF SOLUTIONS OBTAINED BY THE PRINCIPLE OF MAXIMUM NONSPECIFICITY

	$\|X\|$	$\|A\|$	$\|B\|$	$\|A \cap B\|$	a	b	$m(X)$	$m(A)$	$m(B)$	$m(A \cap B)$
$K_1 < 1$	10	5	5	2	.7	.5	0	.5	.3	.2
$K_1 = 1$	10	5	4	2	.8	.6	[0, .2]	[.2, .4]	[0, .2]	[.4, .6]
$K_1 > 1$	20	10	12	4	.4	.5	.5	0	.1	.4

Observe that, due to the linearity of the measure of nonspecificity, the use of the principle of maximum nonspecificity leads to linear programming problems. This is a great advantage when compared with the maximum entropy principle. The use of the latter leads to nonlinear optimization problems, which are considerably more difficult computationally.

The use of the maximum nonspecificity principle does not always result in a unique solution, as demonstrated by the case of $K_1 = 1$ in (9.68). If the solution is not unique, this is a good reason to utilize the second type of uncertainty—strife. We may either add the measure of strife, given by (9.54), as a second objective function in the optimization problem, or use the measure of total uncertainty, given by (9.59), as a more refined objective function. These variations of the maximum uncertainty principle have yet to be developed.

Principle of Uncertainty Invariance

Our repertory of mathematical theories by which we can characterize and deal with situations under uncertainty is already quite respectable, and it is likely that additional theories will be added to it in the future. The theories differ from one another in their meaningful interpretations, generality, computational complexity, robustness, and other aspects. Furthermore, different theories may be appropriate at different stages of a problem-solving process or for different purposes at the same stage. Based on various studies, it is increasingly recognized that none of the theories is superior in all respects and for all purposes. In order to opportunistically utilize the advantages of the various theories of we need the capability of moving from one theory to another as appropriate. ... of transformations from one theory to another should be based on some ... principle. When well-established measures of uncertainty are available in the theories involved, the following principle, called a *principle of uncertainty invariance*, is ... to this purpose.

To transform the representation of a problem-solving situation in one theory, T_1, into an equivalent representation in another theory, T_2, the principle of uncertainty invariance requires that:

1. the amount of uncertainty associated with the situation be preserved when we move from T_1 into T_2; and
2. the degrees of belief in T_1 be converted to their counterparts in T_2 by an appropriate scale, at least ordinal.

Requirement 1 guarantees that no uncertainty is added or eliminated solely by changing the mathematical theory by which a particular phenomenon is formalized. If the amount of uncertainty were not preserved, then either some information not supported by the evidence would unwittingly be added by the transformation (information bias), or some useful information contained in the evidence would unwittingly be eliminated (information waste). In either case, the model obtained by the transformation could hardly be viewed as equivalent to its original.

Requirement 2 guarantees that certain properties, which are considered essential in a given context (such as ordering or proportionality of relevant values), are preserved under the transformation. Transformations under which certain properties of a numerical variable remain invariant are known in the theory of measurement as scales.

Due to the unique connection between uncertainty and information, the principle of uncertainty invariance can also be conceived as a *principle of information invariance or information preservation*. Indeed, each model of a decision-making situation, formalized in some mathematical theory, contains information of some type and some amount. The amount is expressed by the difference between the maximum possible uncertainty associated with the set of alternatives postulated in the situation and the actual uncertainty of the model. When we approximate one model with another one, formalized in terms of a different mathematical theory, this basically means that we want to replace one type of information with an equal amount of information of another type. That is, we want to convert information from one type to another while, at the same time, preserving its amount. This expresses the spirit of the principle of information invariance or preservation: no information should be added or eliminated solely by converting one type of information to another. It seems reasonable to compare this principle, in a metaphoric way, with the principle of energy preservation in physics.

Let us illustrate the principle of uncertainty invariance by describing its use for formalizing transformations between probabilistic and possibilistic conceptualizations of uncertainty. The general idea is illustrated in Fig. 9.8, where only nonzero components of the probability distribution $\mathbf{p}$ and the possibility distribution $\mathbf{r}$ are listed. It is also assumed that the corresponding components of the distributions are ordered in the same way: $p_i \geq p_{i+1}$ and $r_i \geq r_{i+1}$ for all $i = 1, 2, \ldots, n-1$. This is equivalent to the assumption that values p_i correspond to values r_i for all $i = 1, 2, \ldots, n$ by some scale, which must be at least ordinal.

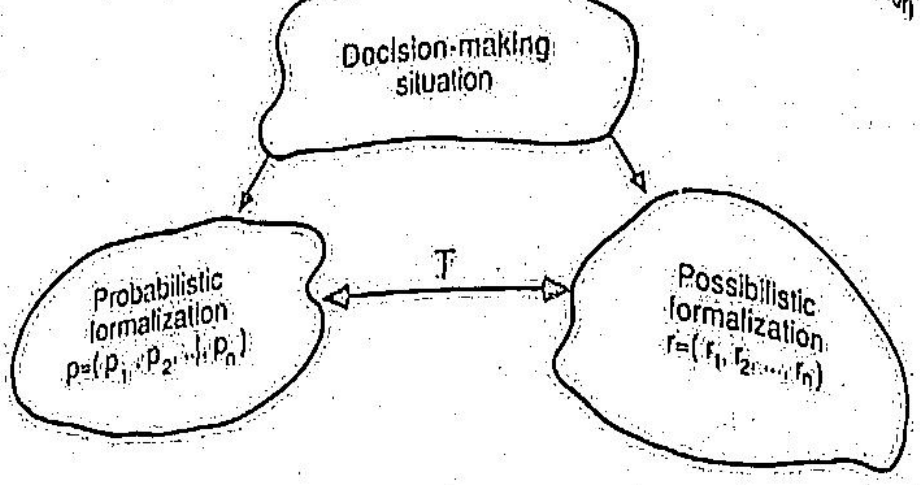


Figure 9.8 Probability-possibility transformations.

Thus far, the following results regarding uncertainty-invariant transformations $\mathbf{p} \leftrightarrow \mathbf{r}$ under different scales have been obtained: (1) transformations based on ratio and difference scales do not have enough flexibility to preserve uncertainty and, consequently, are not applicable; (2) for interval scales, uncertainty-invariant transformations $\mathbf{p} \rightarrow \mathbf{r}$ exist and are unique for all probability distributions, while the inverse transformations $\mathbf{r} \rightarrow \mathbf{p}$ that preserve uncertainty exist (and are unique) only for some possibility distributions; (3) for log-interval scales, uncertainty-invariant transformations exist and are unique in both directions; and (4) ordinal-scale transformations that preserve uncertainty always exist in both directions, but, in general, are not unique.

The log-interval scale is thus the strongest scale under which the uncertainty invariant transformations $\mathbf{p} \leftrightarrow \mathbf{r}$ always exist and are unique. A scheme of these transformations is shown in Fig. 9.9. First, a transformation coefficient α is determined by Equation II, which expresses the required equality of the two amounts of total uncertainty; then, the obtained value of α is substituted to the transformation formulas (Equation I for $\mathbf{p} \rightarrow \mathbf{r}$ and Equation III for $\mathbf{r} \rightarrow \mathbf{p}$). It is known that $0 < \alpha < 1$, which implies that the possibility-probability consistency condition ($r_i \geq p_i$ for all $i = 1, 2, \ldots, n$), is always satisfied by these transformations. When the transformations are simplified by excluding $S(\mathbf{r})$ in Equation II, which for large n is negligible, their basic properties (existence, uniqueness, consistency) remain intact.

For ordinal scales, uncertainty-invariant transformations $\mathbf{p} \leftrightarrow \mathbf{r}$ are not unique. They result, in general, in closed convex sets of probability or possibility distributions, which are obtained by solving appropriate linear inequalities constrained by the requirements of normalization and uncertainty invariance. From one point of view, the lack of uniqueness is a disadvantage of ordinal-scale transformations. From another point of view, it is an advantage since it allows us to impose additional requirements on the transformations. These additional requirements may be expressed, for example, in terms of second-order properties, such as projections, noninteraction, or conditioning.

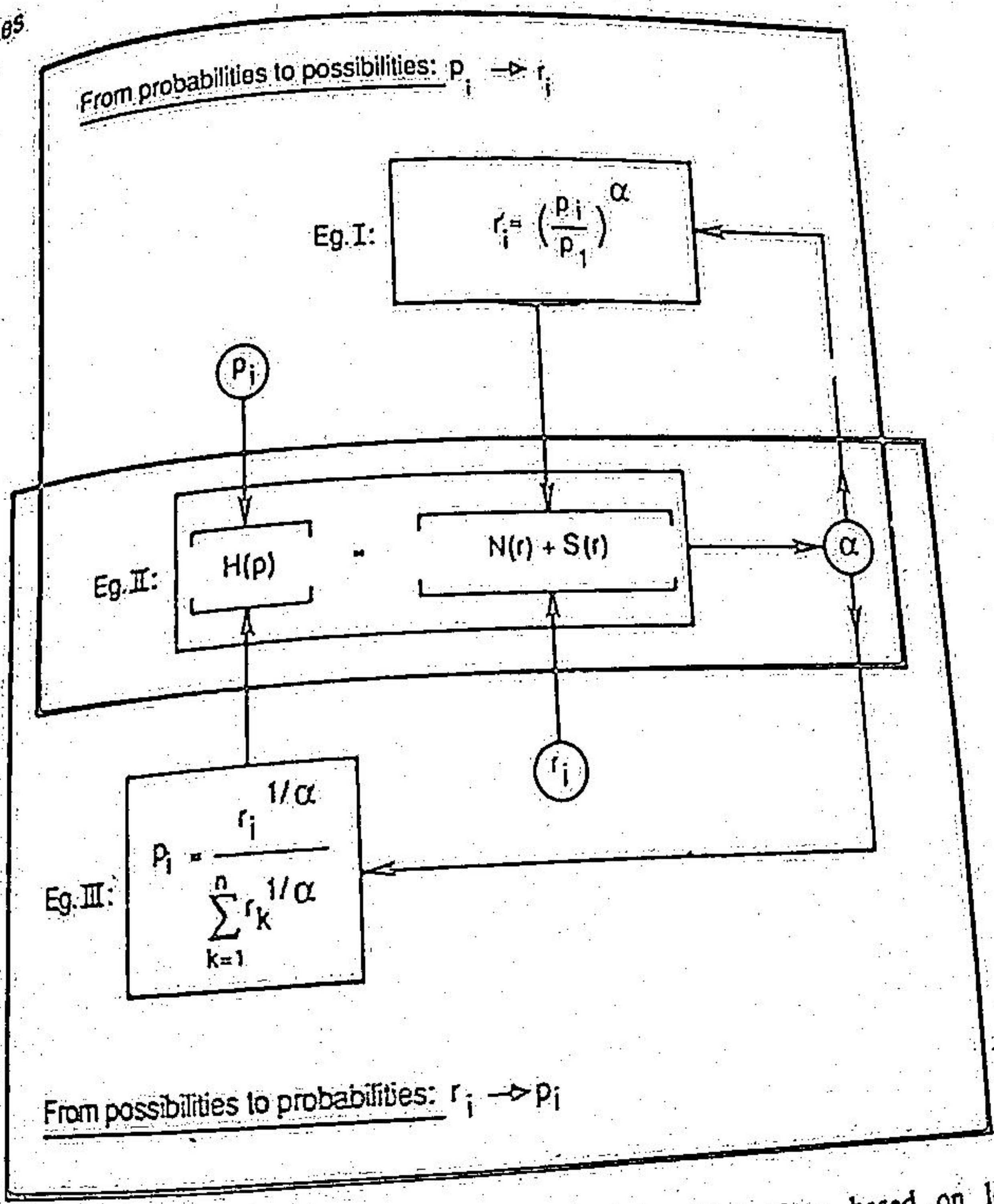


Figure 9.9 Uncertainty-invariant probability-possibility transformations based on log-interval scales.