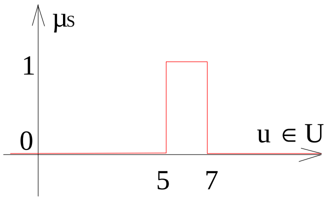
➢ 28th September 1999

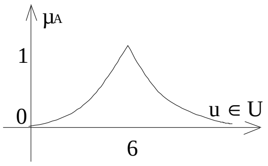
Introduction

(1) **Crisp** Set. Let U = {u_1,...,u_n}, where U is the universal/ordinary set. P(U) is the power set of U, the class of all *subsets*. |U| = n is the cardinality of U. S ⊂ U is the characteristic function, μ_S: U→ {0,1}. S $\mu_S(u_j)$ = 0 for u_j ∈ U, u_j ∉ S; and 1 for u_j ∈ U, u_j ∈ S. Example: U = **R**; ≤ S ≤ 7. The crisp set is as shown in the *diagram*.

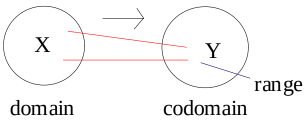


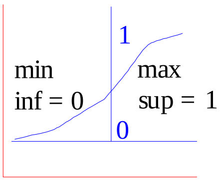
(2) **Fuzzy** Sets. A fuzzy set A defined on U is *characterised* by its membership function μ_A: U → [0,1]. $\mu_A(u_j)$ is the *degree* of membership. An empty fuzzy set A is **defined** as $\mu_A(u_j)$ = 0 for all u_j ∈ U. The *Universal* (fuzzy) set A is defined as $\mu_A(u_j)$ = 1 for all u_j ∈ U. The **Support** of the fuzzy set A on U is the crisp set defined as **supp A** = {u_j | $\mu_A(u_j)$ > 0}. The **height** of the fuzzy set A on U is the height, or hgt A = $\max_{u_j \in U} \mu_A(u_j)$. A fuzzy set A on U is called **normal** iff height A = 1, otherwise it is called *subnormal*.

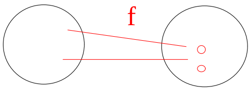
Example (U is finite): U = {Ford, Vauxhall, Fiat); A = Gareth's car; A = {(Fo, 0.4), (V, 0.1), (Fi, 0.5)}. So Supp A = U; height A = 0.5 — so A i *subnormal*. Fuzzy Numbers. Let U = **R** and A = "about 6". $\mu_A(u)$: $\mu_A(6$ should be 1; $\mu_A(u)$ should be *symmetrical* about 6; and $\mu_A(u)$ should decrease as u goes **away** from 6. Should get a graph as shown, but could have **any** pattern.



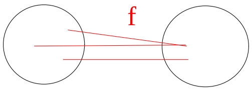
➢ 30th September 1999



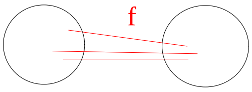




one-to-one, injective



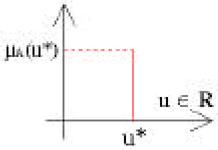
surjective



bijective, one-to-one correspondence

Example: U = {u_1,...,u_6}; A = "Red pencil"; A = {(u_1, 0.1) (u_2, 0.2) (u_3, 0.8) (u_4, 0.3) (u_5, 0.7) (u_6, 0.5)}. Here, support A = U; height A = 0.8, and core A = ϕ. The **core** of a fuzzy set A on U is a crisp set: CoreA = {u_j | u_j ∈ U, $\mu_A(u_j)$ = 1}. A *singleton* is a fuzzy set whose <u>support</u> has cardinality 1.

Example: U = {u_1,...,u_6}; B = {(u_3, 0.8)} (singleton). So B = {(u_1,0) (u_2,0) (u_3,0.8 (u_4,0) (u_5,0) (u_6,0)}. Singleton sets can b *normal* or *subnormal*.

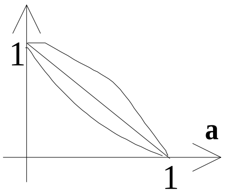


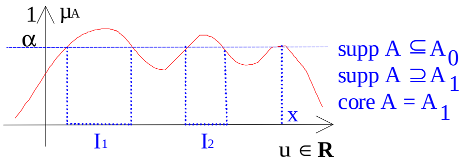
Cardinality and Complement of a Fuzzy Set

Let U = {u_1,...,u_n}, and let A = a *fuzzy* set on U, with |U| = u. The **cardinality** of a fuzzy set A on U is |A| = $\Sigma^{n}_{i=1}$ $\mu_A(u_j)$. Example: Let U = {u_1,...,u_6}; A = {u_3,u_4,u_6}; so A = {(u_1,0) (u_2,0) (u_3,1) (u_4,1) (u_5,0) (u_6,1)}. Now $\Sigma^{i=1}_{6}$ $\mu_A(u_i)$ = 3 = |A|. The relative **cardinality** of a fuzzy set A on U is ||A|| = $^{|A|}/_{|U|}$ = (in this example) = $^{3}/_{6}$ = ½.

Let U = {u_1,...,u_n}, and let S ⊂ U. (S is a **crisp** set). $\bar{S}$ is the *complement*, where $S \cap \bar{S}$ = ϕ (the *non-contradiction* principle) and $S \cup \bar{S}$ = U (the law of *excluded* middle). Now let A be a fuzzy set on U, and define h: [0,1]→ [0,1], with h(a)∈ [0,1], by h(a) = $\mu_A(u_j)$: u_j ∈ U.

(1) h is a function of *1 argument only*. (2) $h(0) = 1$, $h(1) = 0$ (the **complement** applies to crisp sets). (3) h is continuous and *strictly monotonically decreasing*. ($a_1 > a_2$ with a_1 and $a_2 \in [0,1] \Rightarrow h(a_1) < h(a_2)$, as shown in the *diagram*). (4) h is **involutive**: $h(h(a)) = a$. $h(a) = 1-a$. The *complement* of a fuzzy set A on U is a fuzzy set $\bar{A}$ on U such that $\mu_{\bar{A}}(u_j) = 1-\mu_A(u_j)$ for all $u_j \in U$. (5) If $a_1, a_2 \in [0,1]$ so that $a_1+a_2 = 1$, then $h(a_1)+h(a_2) = 1$.





λ-complement (**Sugeno**). $\bar{A}^{\lambda}$, the λ-complement of A on U, is a *fuzzy* set such that $\mu_{\bar{A}^{\lambda}}(u_j) = \frac{1-\mu_A(u_j)}{1+\lambda\mu_A(u_j)}$, where $\lambda \in (-1, \infty)$. (Note: an **OPEN** interval). This satisfies *properties* 1-4. Definition: α-cut or α-level set. Let $\alpha \in [0,1]$. Let A be a **fuzzy** set on U. The α-cut (or α level set) of A is a **crisp** set defined as $A_\alpha = \{u_j \mid \mu_A(u) \geq \alpha\}$. In the **diagram**, $A_\alpha = I_1 \cup I_2 \cup \{x\}$.

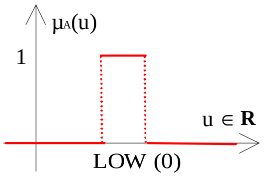
➢ 1st October 1999

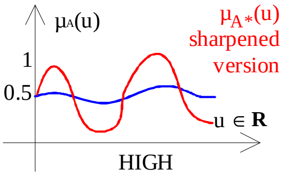
α-cut: *decomposition* (representation) Theorem: Let A be a fuzzy set on U. We can say that $\mu_A(u_j) = \sup_{\alpha\in[0,1]} \min(\alpha, \mu_{A_\alpha}(u_j))$. **Example**: Let $U = \{u_1, \ldots, u_5\}$, where $U = \{(u_1, 0.6)\ (u_2, 0.3)\ (u_3, 0.1)\ (u_4, 0.7)\ (u_5, 0.6)\}$. Now consider α. For $\alpha \leq 0.6$, $\mu_{A\alpha}(u_1) = 1$. For $\alpha > 0.6$, $\mu_{A\alpha}(u_1) = 0$. Consider the *minimum*, $\min(\alpha, \mu_{A\alpha}(u_1))$. This is α for $\alpha \leq 0.6$, and **zero** for $\alpha > 0.6$.

Distances between *fuzzy* sets. Define $d: U\times U \rightarrow \mathbf{R}^+$ as the distance. *Properties*: (1) Non-negative: $d(x,y) \geq 0$. ($x = y \Rightarrow d(x,y) = 0$). (2) **Symmetry:** $d(x,y) = d(y,x)$. (3) Triangle rule: $d(x,z) \leq d(x,y) \underline{*} d(y,z)$, where $\underline{*}$ is an operation *associated* with α. **Euclidean** distance: +.

μ_j	u_1	u_2	u_3	u_4	u_5
μ_{B_1}	1	1	0	1	0
μ_{B_2}	0	1	1	1	0

Consider $B_1, B_2 \subset U$ (crisp subsets). The Hamming distance in the table is 2 (the number of disagreements). Let A and B be **fuzzy** sets on U. The *Hamming* distance between A and B is defined as $d_H(A,B) = \Sigma_{i=1}^{n} |\mu_A(u_i)-\mu_B(u_i)|$. The **Relative** Hamming Distance is defined as $\delta_H = {}^{1}/_{n} d_H(A,B)$ $[|U| = n]$. Here, $0 \leq \delta_H \leq 1$. The *Euclidean* Distance between A and B is defined as $d_E(A,B) = \sqrt{(\Sigma^{n}_{i=1}(\mu_A(u_i)-\mu_B(u_i))^2)}$. The Relative Euclidean *Distance* is $\delta_E = {}^{1}/_{\sqrt{(n)}}\, d_E(A,B)$. Remember that $0 \leq \delta_E \leq 1$.



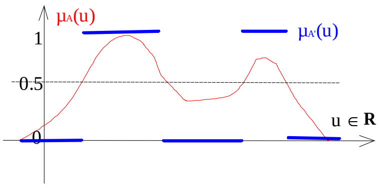


Measures of Fuzziness. Definition: A *sharpened* version of a fuzzy set A on U is any fuzzy set A* such that $\mu_{A^*}(u_j) < \mu_A(u_j)$ if $\mu_A(u_j) < 0.5$; $\mu_{A^*}(u_j) > \mu_A(u_j)$ if $\mu_A(u_j) > 0.5$; and $\mu_{A^*}(u_j) = \mu_A(u_j)$ if $\mu_A(u_j) = 0.5$.

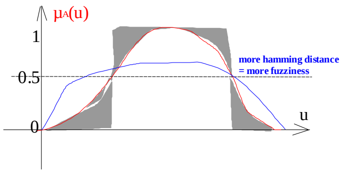
Properties of a measure of fuzziness. Let $\tilde{P}(U)$ be the class of *all fuzzy sets* on U. A **measure** of fuzziness, $H: \tilde{P}(U) \rightarrow \mathbf{R}^+$, is defined by: P1: *Sharpness:* H(A) is a minimum iff $\mu_A(u_i) = 0$ or 1 for all $u_i \in U$. P2: *Maximality:* H(A) is a maximum iff $\mu_A(u_i) = 0.5$ for all $u_i \in U$. P3: *Resolution:* $H(A) \geq H(A^*)$, where $H(A^*)$ is a *sharpened* version of A. P4: *Symmetry:* $H(A) = H(\bar{A})$, where $\bar{A}$ is the complement.

Entropy based measure of fuzziness (Deluca and Terminy): $H_E(A) = -k\Sigma^{n}_{i=1}(\mu_A(u_i)\log(\mu_A(u_i)) + \mu_{\bar{A}}(u_i)\log(\mu_{\bar{A}}(u_i)))$. (k is a *constant).* Another measure: $-k\Sigma^{n}_{i=1}(\mu_A(u_i)\log(\mu_A(u_i)) + (1-\mu_A(u_i))(\log(1-\mu_A(u_i)))$.

➢ 5th October 1999



Definition: the **closest** crisp set A' to a fuzzy set A on U is *defined* as $\mu_{A'}(u_j) = 0$ if $\mu_A(u_j) \le 0.5$, and 1 if $\mu_A(u_j) > 0.5$. Example: U $= \{u_1,u_2,u_3\}$; A $= \{(u_1,0.3), (u_2,0.5)\ (u_3,0.7)\}$. So A' $= \{(u_1,0)\ (u_2,0$ $(u_3,1)\}$, or A' $= \{u_3\}$.



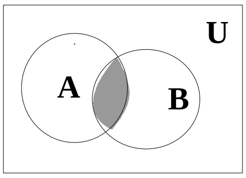
Question: is the closest crisp set a *sharpened* version of A? **Answer:** No! — 0.5 is the problem. Kaufmann — the <u>linear</u> index of fuzziness: (a measure of fuzziness): $\upsilon(A) = {}^2/_n\, d_H(A,A')$. ($d_H$ is the *Hamming* Distance; $n = |U|$). The **quadratic** index is $\eta(A) = {}^2/_{\sqrt{n}}\, d_E(A,A')$.

Definition: Two fuzzy sets A & B on U are *equivalent* iff $\mu_A(u_j) = \mu_B(u_j)$ for all $u_j \in U$.
Definition: A fuzzy set A on U is **included** in a fuzzy set B on U iff $\mu_A(u_j) \le \mu_B(u_j)$ for all $u_j \in U$.
Definition: A fuzzy set A is *strictly included* on B iff $\mu_A(u_j) < \mu_B(u_j)$ for all $u_j \in U$.

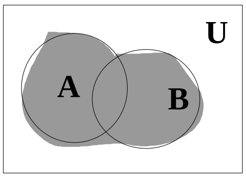
The **extension** principle: Let A be a fuzzy set on U, and let f map U onto V, f: U→ V. Then a fuzzy set B is induced on V, **defined** as $\mu_B(v) = \max_{u:f(u)=v} \mu_A(u)$. Remember that $v \in V$, and if there is a v such that there *exists* $u \notin U$ for which $v = f(u)$, then $\mu_B(v) = 0$. Example: U = {a,b,c,d}; V = {x,y,z}. Looking at the *table*, (**not** injective or surjective or bijective), we see that if we let A = {(a,0.5) (b,0.3) (c,0.7) (d,0.3)}, then B = {(x,0.7) (y,0.5) (z,0)}.

u	a	b	c	d
f(u)	y	x	x	y

<u>**Laws**</u>. $A\cup(B\cup C) = (A\cup B)\cup C$; $A\cap(B\cap C) = (A\cap B)\cap C$ are *associativity*. $A\cup B = B\cup A$ and $A\cap B = B\cap A$ are *commutativity*. $A\cap A = A$ and $A\cup A = A$ are *idempotence*. $A\cap U = A$ and $A\cup\phi = A$ are *identity*. Note: U is an **identity** with respect to intersection, and ϕ is an **identity** with respect to union.



A∩B



A∪B

$A\cup(B\cap C) = (A\cup B)\cap(A\cup C)$ and $A\cap(B\cup C) = (A\cap B)\cup(A\cap C)$ are *distributivity*. $(A\cup B)^c = A^c\cap B^c$ and $(A\cap B)^c = A^c\cup B^c$ are *De Morgan's laws*. $(A^c)^c = A$ is *double negation* (involution). $A\cup A^c = U$ is the *law of excluded middle*. $A\cap A^c = \phi$ is the *non-contradiction principle*. Example: $(A\cap B)^c\cup(A^c\cap B^c\cap C)^c\cup A = U$ — prove this. LHS: $= (A^c\cup B^c)\cup(\ldots)\cup A = (A^c\cup A)\cup B^c\cup(\ldots) = U\cup B^c\cup(\ldots) = U\cup D = U$. **Proved**.

➢ 6th October 1999

<u>Tutorial</u>

Q: Define, using their **membership** functions, a *crisp* set and a *fuzzy* set. What is the <u>difference</u> between the two? A: A **crisp** set S is a subset of U having characteristic function μ_S mapping u onto {0,1}. $\mu_S(u_j) = 0$ if $u_j \in U$, $u_j \notin S$; $\mu_S(u_j) = 1$ if $u_j \in U$, $u_j \in S$.

A **fuzzy** set A defined on U is characterised by its *membership* function, μ_A: U→[0,1]. $\mu_A(u_j)$ is the *degree* of membership. The difference between a *crisp* and *fuzzy* set is that a <u>**membership function**</u> of a crisp set must have values 0 or 1; and a <u>**membership function**</u> of a fuzzy set can have *any* value in the interval [0,1].

Q: Design a fuzzy set "*slow*" (assuming $u \in \mathbf{R}^+$ is m.p.h.). Is this a *normal* fuzzy set? Explain your answer. A: Assume a **finite** number of speeds. Let $A \subset \mathbf{R}^+$ be a car's speed: A = {10, 30, 50, 70, 100}. Let A = (10, 0.9), (30, 0.7), (50, 0.5), (70, 0.2), (100, 0.1). This is *not* a normal fuzzy set because the height A is not 1. In fact, it is 0.9.

u	1	2	3	4	5	6	7	8
$\mu_A(u)$	0.3	0.9	0.4	1	0.7	0	0.6	0.7
$\mu_B(u)$	1	0.2	0.5	0.7	0.3	0.9	0.6	1

Q: Let A and B be *fuzzy* sets on U defined by the table shown. (a) Find the **core** and the **support** of each set. (b) Find the 0.5 cut of the two sets. (c) Calculate the **cardinalities** and relative cardinalities of A and B. (d) Calculate the *linear measure of fuzziness* $\nu(A)$ of A. A: Core A = $\{u_4\}$; Core B = $\{u_1,u_8\}$; supp A = U-$\{u_6\}$; supp B = U. (b) $A_{0.5} = \{u_2,u_4,u_5,u_7,u_8\}$. $B_{0.5} = \{u_1,u_3,u_4,u_6,u_7,u_8\}$.

(c) $|A| = \Sigma_{i=1}^{8}\ \mu_A(u_j) = 0.3+0.9+\ldots+0.7 = 4.6$. $\|A\| = {}^{|A|}/_{|U|} = {}^{4.6}/_{8} = 0.575$. $|B| = 1+0.2+\ldots+0.6+1 = 5.2$. $\|B\| = {}^{5.2}/_{8} = 0.65$. (d) $\nu(A) = {}^{2}/_{n}d_H(A,A')$, $n = |U|$, d_H = *Hamming* distance. So $A' = \{(u_1,0)\ (u_2,1)\ (u_3,0)\ (u_4,1)\ (u_5,1)\ (u_6,0)\ (u_7,1)\ (u_8,1)\}$. The *Hamming* distance is $\Sigma_{i=1}^{n}|\mu_A(u_i)-\mu_{A'}(u_i)| = |0.3-0| + |0.9-1| + |0.4-0| + |1-1| + |0.7-1| + |0-0| + |0.6-1| + |0.7-1| = 1.8$. So $\nu(A) = {}^{2}/_{8}\times 1.8 = 0.45$.

u	1	2	3	4	5	6
$\mu_A(u)$	0.3	0.8	0.9	0.6	1	0.4
$\mu_{\bar{A}}(u)$	0.7	0.2	0.1	0.4	0.9	0.6
$\mu_{A^*}(u)$	0.2	0.9	1	0.7	0	0.3
$\mu_{A'}(u)$	0	1	1	1	0	0

Q: **Let** $U = \{u_1,\ldots,u_6\}$, and let A be a *fuzzy* set on U: $A = \{u_1/0.3,\ u_2/0.8,\ u_3/0.9,\ u_4/0.6,\ u_5/0.1,\ u_6/0.4\}$ Give in a *table the following:* $\mu_A(u)$; $\mu_{\bar{A}}(u)$; $\mu_{A^*}(u)$; and $\mu_{A'}(u)$. A: as shown.

Q: If $A_\alpha = \{u \mid u \in U, \mu_A(u) \geq \alpha\}$, show *that* $\alpha_1 \geq \alpha_2 \Leftrightarrow A_{\alpha_1} \subseteq A_{\alpha_2}$. A: **Proof**: (1) $\alpha_1 \geq \alpha_2 \Rightarrow A_{\alpha_1} \subseteq A_{\alpha_2}$. Solution by **definition:** (1) $A_{\alpha_1} = \{u \mid \mu_A(u) \geq \alpha_1\}$ for $u \in A_{\alpha_1} \Rightarrow \mu_A(u) \geq \alpha_1$. But **since** $\alpha_1 \geq \alpha_2$, (given), this *implies that* $\mu_A(u) \geq \alpha_1 \geq \alpha_2$, implying that $\mu_A(u) \geq \alpha_2$. So $A_{\alpha_1} \subseteq A_{\alpha_2}$.

(2) **Prove** that $A_{\alpha_1} \subseteq A_{\alpha_2} \Rightarrow \alpha_1 \geq \alpha_2$. *Since* $A_{\alpha_1} \subseteq A_{\alpha_2}$, then there exists a $u \in A_{\alpha_2}$, but $u \notin A_{\alpha_1}$. Then, by **definition,** $\mu_A(u) \geq \alpha_2$, and $\mu_A(u) < \alpha_1$. This *implies that* $\alpha_1 > \mu_A(u) \geq \alpha_2$, implying that $\alpha_1 \geq \alpha_2$.

➢ 12th October 1999

Let A and B be *fuzzy* sets on U_1 and U_2 respectively. A **cartesian product** of A and B is a fuzzy set A×B such that $\mu_{A\times B}(u_1,u_2) = \min\{\mu_A(u_1), \mu_B(u_2)\}$ (for $u_1 \in U_1$, $u_2 \in U_2$). Example: $U_1 = \{1,2,4,6\}$; $U_2 = \{a,d\}$; A = {(1,0.3) (2,0) (4,0.1) (6,0.7)}; B = {(a,0.1) (d,0.6)}. So $U_1 \times U_2$ = {(1,a) (2,a) (4,a),...,(6,d)}, and $U_{A\times B}(a,1) = 0.1$. The m^{th} **power** of a fuzzy set A on U is a fuzzy set *denoted by* A^m, such that $\mu_{A^m}(u) = (\mu_A(u))^m$ for all $u \in U$. When m = 2, the *operation* is called "**concentration**". It "shrinks" the *membership* function.

Propositions and Logic Functions

Proposition is a statement which is *true* or *false*. Negation and the proposition have opposite truth values. Let v be a logic variable, so $v \in \{T,F\}$ or $v \in \{1,0\}$. A function of **one** or more logic variables is defined as $f: \{T,F\}^n \rightarrow \{T,F\}$. (LHS = *Cartesian* product). This is called a **logic** function. Logic operators: **conjunction** ($\wedge$), **disjunction** ($\vee$), **negation** (bar, ¯).

Logic functions can be *expressed* in two ways: as **truth tables** and as **logic expressions.** Use truth tables to *define* logic operations. Assume that a and b are logic variables. Conjunction is “**And**”; disjunction is “**Or**”; and negation is “**Not**”. Values are as shown in the table.

a	b	$a \wedge b$	$a \vee b$	$\bar{a}$
T	T	T	T	F
T	F	F	T	F
F	T	F	T	T
F	F	F	F	T

Logic *expressions* are defined (recursively) as: (1) <u>True</u> and <u>False</u> are logic expressions; (2) If a is a logic variable, then both a and $\bar{a}$ are logic expressions; (3) If a and b are logic expressions, then $a \wedge b$ and $a \vee b$ are logic expressions too; (4) The **only** logic expressions are the ones specified by steps 1-3. Examples: F, $x \wedge F$, $x_1 \wedge \bar{x}_2 \vee x_3$.

We can “*bend the rules*” and come up with: (5) If a is a logic expression, then $\bar{a}$ is a logic expression too; (6) If a and b are logic expressions, then $a \Rightarrow b$ is a logic expression too. Implication. “If a then b”; “a involves b”; “from a follows b”. **Example**: if I EAT then I GET FAT. “I EAT” is the *premise,* and “I GET FAT” is the *consequent*.

Two logic functions are equivalent or identical if they yield the **same** values for all values of their arguments. **Tautology** — if a logic function gives value T for all values of its arguments. **Contradiction** — gives values F for all values of its arguments. Here, $(a \wedge (a \Rightarrow b)) \Rightarrow b$ *(modus ponens);* $(\bar{b} \wedge (a \Rightarrow b)) \Rightarrow \bar{a}$ *(modus talbus);* $((a \Rightarrow b) \wedge (b \Rightarrow c)) \Rightarrow (a \Rightarrow c)$ *(syllogisms rule)*. In the **second** table, because the last column consists of T’s, it is a *tautology*.

a	b	$a \Rightarrow b$	$\bar{a}$	$\bar{a} \vee b$
T	T	T	F	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

a	b	$a \Rightarrow b$	$a \wedge (a \Rightarrow b)$	$(a \wedge (a \Rightarrow b)) \Rightarrow b$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

<u>Tutorial</u>

Q: Let A and B be *fuzzy* sets in U. Prove or disprove the following: (a) $H(A) > H(B) \Rightarrow |A| > |B|$; (b) $|A| = |B| \Rightarrow H(A) = H(B)$; (c) $A = B \Rightarrow H(\bar{A}) = H(B)$. (H is the *measure* of fuzziness.) A: (a) Disprove by *example*. Let $|A| = \Sigma_{i=1}^{n} \mu_A(u_i)$; $U = \{u\}$, and let $A = \{(u,0.3)\}$. So $|A| = 0.3$; $A' = \phi = \{(u,0)\}$; $|\mu_A(u)-\mu_{A'}|(u)| = 0.3$; and $\nu(A) = {}^{2}/_{1} \times 0.3 = 0.6$.

Now let $B = \{(u,0.9)\}$. So $|B| = 0.9$; $B' = u = \{(u,1)\}$; $|\mu_B(u)-\mu_{B'}(u)| = 0.1$; and $\nu(A) = 0.2$. **Here,** $|B| > |A|$, and $\nu(A) > \nu(B)$. This *does not match* what is <u>given</u> in the question. (b) Let $U = \{u_1,u_2\}$, and let $A = \{(u_1,1)\ (u_2,0)\}$. So $|A| = 1$; $A' = \{(u_1,1)\ (u_2,0)\}$; $d_H = |(1-1)+(0-0)| = 0$; so $\nu(A) = 0$. Now let $B = \{(u_1,0.5)\ (u2,0.5)\}$. So $|B| = 1$; $B' = \{(u_1,0)\ (u_2,0)\}$; $d_H = 1$; so $\nu(B) = 1$. **Different**! (c) Let A = B, then $\mu_A(u) = \mu_B(u)$ for all $u \in U$. Then $H(A) = H(B)$. From the property of *symmetry,* $H(A) = H(\bar{A})$*, therefore* $H(\bar{A}) = H(B)$.

Q: Prove that $\nu(A)$ satisfies: (a) the *sharpness* property: $\nu(A) = \min$ iff A is a crisp set; (b) the maximality property: $\nu(A) = \max$ iff A is such that $\mu_A(u) = 0.5$ for all $u \in U$. Q: **Show** that the λ-complement ($h(a) = {}^{1-a}/_{1+\lambda a}$; $\lambda \in (-1, \infty)$) is (a) *monotone decreasing* for $a \in [0,1]$; (b) *involute*.

(a) $\nu(A) = {}^{2}/_{n} d(A,A') = {}^{2}/_{n} \Sigma_{i=1}^{n} |\mu_A(u_i)-\mu_{A'}(u_i)|$. This is a *summation* of +ve terms. Therefore, it reaches a **minimum** iff each term is zero, i.e. $|\mu_A(u_i)-\mu_{A'}(u_i)| = 0$, i.e. $\mu_A(u_i) = \mu_{A'}(u_i)$ for all $u_i \in U$. $\mu_{A'}(u_i) \in \{0,1\}$ by *definition* (A' is a **crisp** set), therefore $\mu_A(u_i) \in \{0,1\}$ too. **Hence** A is a *crisp* set.

(b) $\nu(A) = {}^{2}/_{n} \Sigma_{i=1}^{n} |\mu_A(u_i)-\mu_{A'}(u_i)|$. This is a **summation** of +ve terms. It takes its *maximum* when each term is **maximal**, i.e. iff each term is *maximal*. Now $\max|\mu_A(u_i)-\mu_{A'}(u_i)| = 0.5$ iff $\mu_A(u_i) = 0.5$ and $\mu_{A'}(u_i) = 0$.

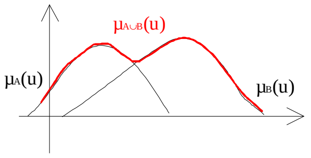
➢ 14th October 1999

(1) $a\wedge b = b\wedge a$; $a\vee b = b\vee a$. (2) $a\wedge(b\wedge c) = (a\wedge b)\wedge c$; $a\vee(b\vee c) = (a\vee b)\vee c$. (3) $a\vee(b\wedge c) = (a\vee b)\wedge(a\vee c)$; $a\wedge(b\vee c) = (a\wedge b)\vee(a\wedge b)$. (4) $a\wedge a = a$; $a\vee a = a$. (5) $a\wedge T = a$; $a\wedge F = F$; $a\vee T = T$; $a\vee F = a$. (6) $a\vee\bar{a} = T$; $a\wedge\bar{a} = F$. (7) $(\overline{(\bar{a})}) = a$. (8) $\overline{(a\wedge b)} = \bar{a}\vee\bar{b}$; $\overline{(a\vee b)} = \bar{a}\wedge\bar{b}$.

Q: Prove **modens tallens** by manipulating *logic* expression (Boolean expressions). A: We have to prove the following: $(\bar{b}\wedge(a \Rightarrow b)) \Rightarrow \bar{a}$. Now $a \Rightarrow b = \bar{a}\vee b$. $(\bar{b}\wedge(\bar{a}\vee b)) \Rightarrow \bar{a} = \overline{(\bar{b}\wedge(\bar{a}\vee b))}\vee\bar{a} = (\overline{(\bar{b})}\vee\overline{(\bar{a}\vee b)}\vee\bar{a} = (\overline{(\bar{b})})\vee\overline{(\bar{a}\vee b)})\vee\bar{a} = (b\vee\overline{(\bar{a})}\wedge\bar{b})\vee\bar{a} = ((b\vee a)\wedge(b\vee\bar{b}))\vee\bar{a}$. Now *because* $b\vee\bar{b} = T$, this becomes $(b\vee a)\vee\bar{a} = b\vee(a\vee\bar{a}) = b\vee T = T$. QED. The **expression** $((\bar{b}\wedge(a\Rightarrow b))\Rightarrow\bar{a}$ can also be proved by a *truth table*, as shown.

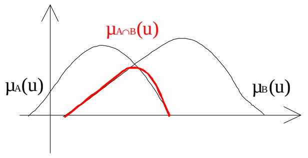
		[1]	[2]	[3]	[4]	
a	b	$a\Rightarrow b$	$\bar{b}$	$(2)\wedge(1)$	$\bar{a}$	$(3)\Rightarrow(4)$
T	T	T	F	F	F	T
T	F	F	T	F	F	T
F	T	T	F	F	T	T
F	F	T	T	T	T	T

Operations on Fuzzy Sets

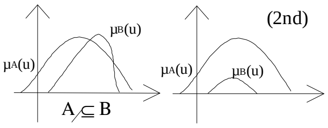


(1) **Union**. Let A & B be *fuzzy sets* on U. The union of A and B is a **fuzzy** set, denoted by $A\cup B$ $(A\vee B)$, with membership *function* $\mu_{A\cup B}(u) = \max\{\mu_A(u), \mu_B(u)\}$ for all $u \in U$. *Example*: $U = \{u_1,u_2,u_3,u_4\}$; $A = \{(0.1, u_1)\ (0.2, u_2)\ (0.6, u_3)\ (0.1, u_4)\}$; and $B = \{(0, u_1)\ (0, u_2)\ (0.3, u_3)\ (0.4, u_4)\}$. So $A\cup B = \{(0.1, u_1)\ (0.2, u_2)\ (0.6, u_3)\ (0.4, u_4)\}$. Now let $C = \{u_1,u_3\}$, and let $D = \{u_3, u_4\}$ (Crisp Sets). It follows that $C\cup D = \{u_1, u_3, u_4\}$. This can be shown in a *table, as shown on the left.*

	u_1	u_2	u_3	u_4
C	1	0	1	0
D	0	0	1	1
$C\cup D$	1	0	1	1



(2) **Intersection**. The intersection of 2 fuzzy sets, e.g. A and B on U, is a fuzzy set, denoted by $A\cap B$ $(A\wedge B)$, with membership *function* $\mu_{A\cap B}(u) = \min\{\mu_A(u), \mu_B(u)\}$ for all $u \in U$. The union and the intersection *correspond* to these for crisp sets.



Properties: (1) $A\cap B \subseteq A\cup B$. *Reminder*: If C and D on U are fuzzy, then $C \subseteq D$ iff $\mu_C(u) \leq \mu_D(u)$ for all $u \in U$; and $C \subset D$ iff $\mu_C(u) < \mu_D(u)$ for all $u \in U$. In the *second* diagram, $B \subseteq A$ and $B \subset A$.

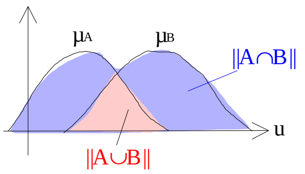
Prove that $A \cap B$ is not strictly *included* in $A \cup B$, i.e. disprove $A \cap B \subset A \cup U$. (By letting A = B). **Proof**: Let $u \in U$ be such that $\mu_A(u) = \mu_B(u)$. Then $\min\{\mu_A(u), \mu_B(u)\} = \mu_{A\cap B}(u) = \mu_{A\cup B}(u) = \max\{\mu_A(u), \mu_B(u)\}$. Therefore, the *strict inequality* is violated, and $A\cap B \subsetneq A\cup B$.

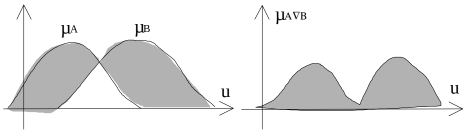
➢ 19th October 1999

max$(a,b) = {}^{a+b+|a-b|}/_{2}$; *min*$(a,b) = {}^{a+b-|a-b|}/_{2}$. (1) **Commutativity:** $A\cup B = B\cup A$ ($\cup$ = max); $A\cap B = B\cap A$ ($\cap$ = min). (2) **Idempotence:** $A\cup A = A$; $A\cap A = A$. (3) **Associativity.** (4) **Distributivity**. (5) $A\cup\phi = A$; $A\cap\phi = \phi$; $A\cup U = U$; $A\cap U = A$. (6) *De Morgan's Laws*: $(A\cup B)^c = A^c\cap B^c$; $(A\cap B)^c = A^c\cup B^c$. Now *prove* that $(A\cup B)^c = A^c\cap B^c$. (A and B are *fuzzy sets on U).* $\mu_{A\cup B} = (u_j)$, with $u_j \in U$. $1-\mu_{A\cup B}(u_j) = \mu_{A\cap B}$ (for all $u_j \in U$) $= 1-\max\{\mu_A(uj), \mu_B(uj)\}$. *Right* hand side: $\min\{1-\mu_A(u_j), 1-\mu_B(u_j)\} = 1-\max\{\mu_A(u_j), \mu_B(u_j)\}$. So LHS = RHS. **Bounded Sum:** Let A and B be fuzzy sets on U. The bounded sum is *denoted by* (and is a fuzzy set) $A\oplus B$, where $\mu_{A\oplus B}(u_j) = \min\{1, \mu_A(u_j)+\mu_B(u_j)\}$. **Bounded Difference:** $A|\text{-}|B$; denoted by $\mu_{A|\text{-}|B}(u_j) = \max\{0, \mu_A(u_j)-\mu_B(u_j)\}$.

The example shown in the **table** is the *symmetric difference,* denoted by $A\nabla B$, where $\mu_{A\nabla B}(u_j) = |\mu_A(u_j)-\mu_B(u_j)|$. Note that $|A\nabla B| = d_H$.

	u_1	u_2	u_3	u_4	u_5
A	0	0.3	0.6	1	0.1
B	0.6	0.1	0.6	0.7	0.3
$A\cup B$	0.6	0.3	0.6	1	0.3
$A\cap B$	0	0.1	0.6	0.7	0.1
$A\oplus B$	0.6	0.4	1	1	0.4
$A\|\text{-}\|B$	0	0.2	0	0.3	0
$A\nabla B$	0.6	0.2	0	0.3	0.2

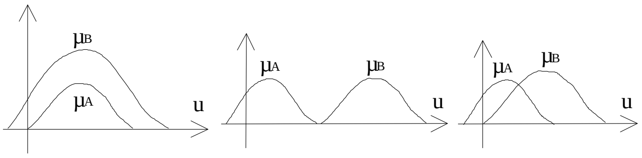


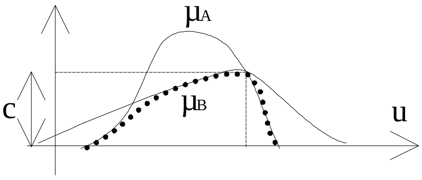


Similarity, Inclusion, Consistency (Measures of Similarity). (1) $S_1(AB) = {}^{\|A\cap B\|}/_{\|A\cup B\|}$. In the *diagram*, $S_1(A,B) \geq 0$. Notes: $S_1(A,B) = 0$ when $A\cap B = \phi$; $S_1(A,B) \leq 1$; $S_1(A,B) = 1$ when A = B. (2) $S_2 = 1-\delta_H(A,B) = 1-|A\nabla B|$. In the **table**, we see *examples* of the use of these measures. Note that $\|A\cap B\| = {}^{1.2}/_3 = 0.4$; $\|A\cup B\| = {}^{2}/_3$; $\|A\nabla B\| = {}^{0.8}/_3$; and so $S_1 = {}^{\|A\cap B\|}/_{\|A\cup B\|} = {}^{1.2}/_2 = 0.6$; and $S_2 = 1-{}^{0.8}/_3 = {}^{11}/_{15}$.

	u_1	u_2	u_3
A	0.3	0.6	1
B	0.4	0.1	0.8
$A\cap B$	0.3	0.1	0.8
$A\cup B$	0.4	0.6	1
$A\nabla B$	0.1	0.5	0.2

Index of inclusion of A in B: $I_1 = {}^{\|A\cap B\|}/_{\|A\|}$. I diagram 1, $I_1 = 1$. In diagram 2, $I_1 = 0$. In diagram 3 $0 \leq I_1 \leq 1$. If we interchange A and B in **diagram** 1, then $I_1 = {}^{\|B\|}/_{\|A\|} \leq 1$.





Index of **Consistency:** Let A and B **fuzzy** sets on U, then $C(A,B) = \text{height}(A\cap B)$. S_1, S_2 and I_1 are integral-type measures. C is a *point-wise* measure.

➢ 21st October 1999

On the next page, we see a table of **fuzzy implications**. The "Y" or "N" indicate whether these fuzzy implications agree with *normal boolean implication.* Think of it as testing values in a truth table, where 1 column contains a; **another** b; and the third $a\Rightarrow b$. Then we test the implications in the table to see if they *match up with the value*s seen in $a\Rightarrow b$. Note: think of values not as "T" or "F", but as "1" and "0". This table is for **REFERENCE ONLY**. Notation: $a = \mu_A(u_j)$, $b = \mu_B(u_j)$.

Mamdani Type	min{a,b}	N	**Zadek**	max{min{a,b}, 1-a}	Y
Larsen	a.b	N	**Standard**	a if a ≤ b, 0 if a > b	Y
Lukasiewicz	min{1, 1-a+b}	Y	**Drastic Product**	a if b = 1, b if a = 1, 0 if a < 1, b < 1.	N
Kleen-Dienes	max{1-a, b}	Y	**Gougen**	1 if a ≤ b, b/a if a > b	Y
Bounded Product	max{0, a+b-1}	N	**Gobelian**	1 if a ≤ b, b if a > b.	Y

Aggregation Operations

Two-place aggregations: A: [0,1]×[0,1]→[0,1], where A(a,b) ∈ [0,1], and a,b ∈ [0,1]. **T-norms** (t-norms) are *intersection* type aggregation operations. **T-co-norms** (t-co-norms) or S-norms (s-norms) are *union* type aggregation operations.

For T-norms, **the following** must all be satisfied: (1) *Limit*: T: [0,1]×[0,1]→[0,1]. (2) *Commutative:* T(a,b) = T(b,a). (3) *Associative:* T(a, T(b,c)) = T(T(a,b), c). (4) *Monotonicity*: for all a, b, c and d such that a > c and b > d, then T(a,b) > T(c,d). (5) One-identity: T(a,1) = T(1,a) = a.

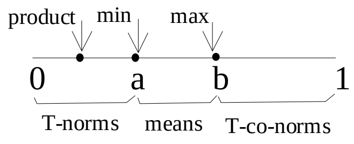
For **T-co-norms**, **the following** must all be satisfied: (1) *Limit*: S: [0,1]×[0,1]→[0,1]. (2) *Commutativity:* S(a,b) = S(b,a). (3) *Monotonicity*: for all a, b, c and d ∈ [0,1] such that a > c, and b > d, then S(a,b) > S(c,d). (4) *Associativity:* S(a, S(b,c)) = S(S(a,b), c). (5) Zero-identity: S(a,0) = S(0,a) = a. [**Idempotence**: T(a,a) = a, or S(a,a) = a. **Archimedean property**: T(a,a) < a, or S(a,a) > a].

The **only** idempotent T-norm is 'min'. The **only** idempotent T co-norm is 'max'. Q: Show that *product* is a T-norm. A: (1) For all a,b ∈ [0,1], a.b ≤ a or a.b ≤ b, and so a.b ∈ [0,1]. (2) **Commutativity:** obvious! (3) **Associativity:** a.(b.c) = (a.b).c. (4) **Monotonicity:** a > c and b > d imply that a.b > c.d. (5) **One-identity**: a.1 = 1.a = a. *All the axioms* are satisfied.

T-norms	**T-co-norms**	**name(s)**
min(a,b)	max(a,b)	min, max
a.b	a+b - a.b	product, probabilistic sum
max{0, a+b-1}	min{1, a+b}	bold union, bounded sum.

➢ 26th October 1999

Duality of t-norms and t-conorms



A **t-norm** T(a,b) and a **t-conorm** S(a,b) are dual with respect to the *standard complement* iff T(a,b) = 1-S(1-a, 1-b), or S(a,b) = 1-T(1-a, 1-b). Also, T(a,b) ≤ min {a,b}, and S(a,b) ≥ max {a,b}. n-place "aggregation operations": A: $[0,1]^n$→[0,1], with $a_1,\ldots,a_n$ (a_i ∈ [0,1]), where A($a_1,\ldots,a_n$) ∈ [0,1], "*representative*".

U	A_1	A_2	A_3	A_4	A
u_1	0.3	0	0.6	0.2	
u_2	0.8	1	0.7	0.6	
u_3	0.4	0.8	0.9	0.9	

From the table, we have different **aggregations**. **Optimistic** aggregation would have column values 0.6, 1 and 0.9 (*max*). **Pessimistic** Aggregation: 0, 0.6, 0.4 (*min*). **Indifferent** Aggregation: 0.3, 0.775, 0.75 (*average*).

Ordered Weighted Averaging Operators (OWA)

Aggr. Operator	OWA $\underline{w}$ $a_1,..,a_n$
max	$[100...0]^T$
min	$[00..001]^T$
average	$[^1/_n\ ^1/_n \ldots\ ^1/_n]^T$
median	$[0...010.,0]^T$ for *odd* n, $[0..^1/_2\ ^1/_2..00]$ for *even* n.
competition jury	$[0\,^1/_{n-2}\,^1/_{n-2}\ldots{}^1/_{n-2}\,0]$

Let $\underline{w} = [w_1,w_2,\ldots,w_n]^T$; and let $a = [a_1,\ldots,a_n]^T$. The *indifferent weighted aggregation is* $\underline{a}^T.\underline{w} = \underline{w}^T.\underline{a}$. $(\Sigma_{i=1}^{n} w_i = 1)$. **Sort** $a_1,\ldots,a_n$ such that $a_{i1} \geq a_{i2} \ldots\ \geq a_{in}$. Example: $\underline{a} = [0.1, 0.6, 0.3, 0.7]^T$; $\underline{w} = [0.3, 0.1, 0.1, 0.5]^T$. The OWA is $[0.7, 0.6, 0.3, 0.1][0.3, 0.1, 0.1, 0.5]^T = 0.21+0.06+0.03+0.05 = 0.35$.

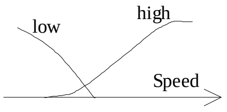
Q: Show that the *probabilistic sum* is a t-conorm. Show its duality with the product. A: Prove the axioms. (1) **Limit** condition: S(a,b) = a+b-a.b; S(a,b) ∈ [0,1] for all a,b ∈ [0,1]. Now a ≥ 1.b.a+b-a.b = (a-ab) + b. Because a-ab ≥ 0 and b ≥ 0, then whole expression is ≥ 0.

(2) **Commutativity:** obvious. (3) **Associativity:** we have to prove that S(S(a,b), c) = S(a, S(b,c)). LHS: (a+b-a.b)+c-(a+b-ab).c = a+**b**+**c**-ab-ac-**bc**+abc = (b+c-bc)+a-ab-ac+abc = (b+c-bc)+a-(b+c-bc).a = S(a,S(b,c)) = RHS. (4) **Zero Identity:** S(a,0) = a+0-0.a = a. (5) **Monotonicity:** *show* that for all a, b, c and d ∈ [0,1] such that a > c and b > d, we have S(a.b) > S(c.d). **Form**: show that S(a.b)-S(c.d) > 0. Now a+b-ab-c-d+cd = (a-c)+(b-d)-ab+cd. Here, (a-c) > 0, (b-d) > 0, and ab > cd. This *implies that* (a-c)+(b-d)-ab+cd = a-c+b-d-ab+cd = a-ab+b-c+cd-d = (a+b)-(c+d)-ab = a+b-c-d-ab+cd = [(a+b)-(c+d)]-[ab-cd]. Both things in **square** brackets are > 0. So we have (a-c)+b(1-a)-**d(a-c)**. *This* is > (a-c)+b(1-a)-b(1-c) ≥ 0.

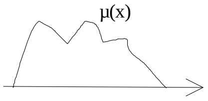
➢ 28th October 1999

Determining Membership Functions

This is done by **Human** (e.g. experts or interviews) or **Automatically** (extract μ's from data). If U is *continuous valued* or *discrete, we have* membership functions. If U is *nominal*, we only have degrees of membership for the elements of U. We can have **Linguistic** Modelling or **Non-Linguistic** Modelling. Difficulties: (1) Expert Specific; (2) Context Specific; (3) Problem Specific. An example of the difficulties is in classifying $\mu_{tall}(183cm)$ = ? This depends on what you are *comparing* to, e.g. males, females, basketball players?

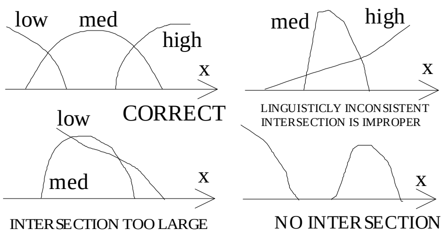


Linguistic Modeling



Non-Linguistic Modeling

Linguistic Modeling



Modelling Methods (Expert/Human)

Method	Question	Comment
Polling (Horizontal)	"Do you think that .. is ..?"	$N_{yes}/N_{total} = \mu(\ldots)$ (need a set of *people*)
Direct Estimation	"What is the degree of membership of...?"	$\mu(\ldots) = (\Sigma_{i=i}^{N_{total}} \mu_i(\ldots))/N_{total}$
Reverse Estimation (Vertical)	"Identify all elements from a given set that you think have degree > 0.3"	0.3 cut. You **need** 1 person
Interval Estimation	"Specify an interval corresponding to your notion of..."	Continuous valued or Discrete U. *Combine* all results as in 1
Pairwise Comparison	"Which of the two members of U is more...?"	We obtain a **preference** relation and calculate μ's from it. |U| should be small.

Assignment 1

Q: **Show** that the λ-complement of fuzzy sets is (i) *monotone decreasing* on [0,1]; (b) *involutive*. A: Let $h(a) = {}^{1-a}/_{a+\lambda a}$ ($\lambda \in (-1, \infty)$). Let $a > b$. Form $h(a)-h(b)$, and show that this is *negative* or *zero*. ${}^{1-a}/_{1+\lambda a} - {}^{1-b}/_{1+\lambda b} = {}^{(1-a)(1+\lambda b)-(1-b)(1+\lambda a)}/_{(1+\lambda a)(1+\lambda b)}$. The *denominator* is always positive for $\lambda > -1$ and $a,b \in [0,1]$. Now $1-a+\lambda b-\lambda ab-1+b-\lambda a+\lambda ab = b(1+\lambda)-a(1+\lambda) = (b-a)(1+\lambda)$. $(b-a)$ is negative by *assumption*, and $(1+\lambda)$ is always > 0. Therefore, $h(a) < h(b)$. (b) $h(h(a))= {}^{1-h(a)}/_{1+\lambda h(a)} = (1-({}^{1-a}/_{1+\lambda a}))/(1+\lambda({}^{1-a}/_{1+\lambda a})) = {}^{1+\lambda a-1+a}/_{1+\lambda a+\lambda-\lambda a} = {}^{(1+\lambda)a}/_{(1+\lambda)} = a$. QED.

Q: Design a *fuzzy* set A on some U, such that $\|A\| = 0.6$; $\nu(A)= 0.6$; and $|A_{0.5}| = 2$. Show and explain the *sequence* of calculations. A: $n = |U| \geq 2$ because $|A_{0.5}| = 2$. Try $n = 3$. Let $a = \mu_A(u_1)$, $b = \mu_A(u_2)$, and $c = \mu_A(u_3)$. For $|A_{0.5}|$ to be 2, we need exactly *2 of the three* to be in [0.5, 1]. It does not matter **which** two we choose, so let $a \geq 0.5$, $b \geq 0.5$, and $c < 0.5$.

$\|A\| = {}^{a+b+c}/_3 = 0.6$. So $a+b+c = 1.8$. Now $\nu(A) = {}^2/_3((1-a)+(a-b)+c) = 0.6$; $2(2-a-b+c) = 1.8$; $a+b-c = 1.1$. Any three numbers that satisfy the *system* of equations, and the 3 inequalities, *is* a solution. So $a+b+c+a+b-c = 2.9$; $a+b = 1.45$; $1.45-c = 1.1$; $c = 0.35$. Pick $a = 0.80$, so $b = 0.65$. Therefore, the set A is $(u_1, 0.80)$, $(u_2, 065)$, $(u_3, 0.35)$. Check: $\|A\| = {}^{0.8+0.65+0.35}/_3 = 0.6$; $\nu(A) = {}^2/_3(0.2+0.35+0.35) = 0.6$.

Q: Prove or **disprove** that (a) $\|A\cap B\| \leq \|A\cup B\|$; (b) $\text{height}(A) = 1 \Leftrightarrow \text{height}(A\cup B) = 1$; (c) if $(\text{core}(A) \neq \phi)$ and $(\text{core}(B) \neq \phi)$, then $\text{height}(A\cap B) = 1$. A: (a) $\|A\cap B\| = {}^1/_n\Sigma^{i=1}{}_n \min(\mu_A(u_i), \mu_B(u_i)) \leq {}^1/_n\Sigma^{i=1}{}_n \max(\mu_A(u_i), \mu_B(u_i)) = \|A\cup B\|$. (b) **height**$(A) = 1 \Leftrightarrow \text{height}(A\cup B) = 1$. *Disprove* by an example: Let $U = \{u\}$; $\mu_A(u) = 0.4$; and $\mu_B(u) = 1$. Then $\mu_{A\cup B}(u) = 1$. But height $\mu_A(u) = 0.4 < 1$. Therefore, *height*$(A\cup B) = 1 \Rightarrow \text{height}(A) = 1$ does not hold, and the **statement** is false. (c) Disprove by an *example*: Let $U = \{u_1,u_2\}$; $A = \{(u_1, 0.1), (u_2, 1)\}$; and $B = \{(u_1,1), (u_2, 0.3)\}$. Then $A\cap B = \{(u_1,0.1)\ (u_2,0.3)\}$; and height $(A\cap B) = 0.3\ (\neq 1)$.

Q: Assuming that the truth table of **implication** is not defined for argument values (F,T) and (F,F), prove, *using truth tables*, that in order for modus ponens to hold, the <u>only</u> choice for the **missing** values is T. A: Denote the missing values by *x and y*. The truth table for modus ponens is as shown. But modens ponens is a <u>tautology</u>, and therefore the last column should contain only T. *Therefore*, x = T, and y = T.

a	b	a⇒b (①)	a∩① (②)	②⇒b
T	T	T	T	T
T	F	F	y	F
F	T	x	x	F
F	F	y	y	F

➢ 2nd November 1999

Automatic Methods

We can have **one** fuzzy set, or “**c**” fuzzy sets. <u>End-point estimation:</u> Let $U = \{u_1,...,u_n\}$, then $U_{max} = \max\{u_i\}$; $U_{min} = \min\{u_i\}$; and $\bar{u} = {}^1/_n\ \Sigma^{i=1}{}_n\ u_i$. Using **Histograms**: see diagram (i). If we have c membership functions, we use *clustering* ($\Sigma_{i=1}{}^c\ \mu_i(u) = 1$ [*shared* degrees of membership]), or the k **nearest** numbers ($\mu_i(u) = {}^{ki}/_k$). An example is shown in **diagram** (ii), where $k = 11$; $\mu_0(u) = {}^6/_{11}$; $\mu_+(u) = {}^4/_{11}$; and $\mu_*(u) = {}^1/_{11}$.

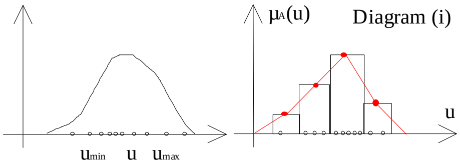


Diagram (i)

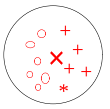
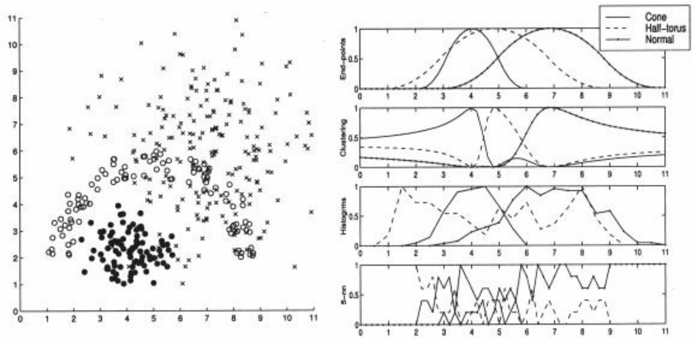


Diagram (ii)



Crisp and Fuzzy Relations

Let X and Y be *crisp* sets. The Cartesian Product is defined by $X \times Y = \{(x,y) \mid x \in X, y \in Y\}$ (an **ordered** pair). If $X \neq Y$, then $X \times Y \neq Y \times X$. If $X = Y$, then $X \times Y = Y \times X$. Example: $X = \{1,2\}$; $Y = \{a,b,c\}$. Then $X \times Y = \{(1,a), (1,b), ..., (2,c)\}$. Let $X_1,...,X_m$ be *crisp* sets. The **Cartesian product** of $X_1,...,X_m$ is a set $X_1 \times X_2 \times ... \times X_m = \{(x_1,x_2,...,x_m) \mid x_1 \in X_1, ..., x_m \in X_m\}$.

A **relation** among crisp sets $X_1,...,X_m$ is a *subset of the Cartesian product* of $X_1 \times X_2 \times ... \times X_m$, so that $R(X_1,X_2,...,X_m) \subset X_1 \times ... \times X_m$. Example: $R(X,Y) = \{(1,b), (2,a)\}$. A *relation* between 2 sets is called a **binary** relation. 3 sets: ternary. 4 sets: quaternary. 5 sets: quinary. In general, we have *n-dimensional* relations.

Example: X_1 = {English, French}; X_2 ={dollar, pound, franc mark}; X_3 = {USA, France, Canada, Britain, Germany}. Therefore, R = *{(English, pound, Britain), (French, franc, France), (English, dollar, USA), (French, dollar, Canada), (English, dollar, Canada)}*. We can represent this in a *matrix* as shown on the right. The matrix shown is for the "**English**" case.

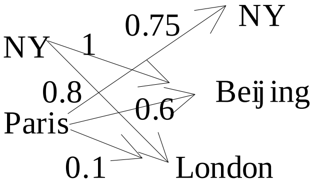
$$\begin{array}{c|ccccc|} & \text{USA} & \text{France} & \text{Canada} & \text{Britain} & \text{Germany} \\ \text{dollar} & 1 & 0 & 1 & 0 & 0 \\ \text{pound} & 0 & 0 & 0 & 1 & 0 \\ \text{franc} & 0 & 0 & 0 & 0 & 0 \\ \text{mark} & 0 & 0 & 0 & 0 & 0 \end{array}$$

➢ 4th November 1999

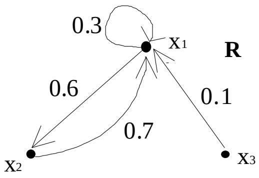
A fuzzy **relation** R on X and Y is a fuzzy set on X×Y. And $\mu_R(x,y) \in [0,1]$. Example: Let X = {NY, Pairs}; Y = {NY, Beijing, London}; and let R = "very far". We could have a *table* as shown. R is a binary fuzzy relation. An n-dimensional *fuzzy* relation is $X_1 \times X_2 \times ... \times X_m$. And $\mu_R(x_1,x_2,...,x_m) \in [0,1]$.

	NY	Beijing	London
NY	0	1	0.8
Paris	0.75	0.6	0.1

Binary Relations (Crisp and Fuzzy)



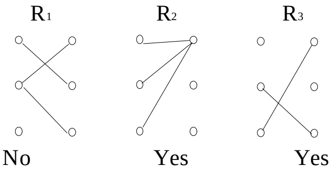
R(X,X) can be expressed as a *directed graph* (a digraph). Consider $X = \{x_1,x_2,x_3\}$ and $R = \{(x_1,x_1,0.3), (x_1,x_2,0.6), (x_2,x_1,0.7), (x_3,x_1,0.1)\}$. Then we will have the *digraph* as shown on the right. But if we wanted to show a **saggital diagram**, we would have a diagram as shown on the *left*.



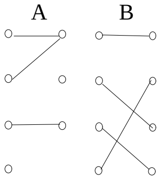
The **domain** of a crisp binary relation R is as follows: dom $R(X,Y) = \{X \mid x \in X, (x,y) \in R\}$. The domain of a *fuzzy binary relation* R is a fuzzy set on X such that $\mu_{\text{dom } R}(x) = \max_{y \in Y} \mu_R(x,y)$. So in our example, $\mu_{\text{dom } R}$(**Paris**) = 0.75.

The *range of a crisp binary relation* R is as follows: ran R(X,Y) = {y | y$\in$Y, (x,y)$\in$R}. The range of a *fuzzy binary relation* R is a fuzzy set on Y such that $\mu_{\text{ran R}}(y) = \max_{x\in X} \mu_R(x,y)$. So in the example, $\mu_{\text{ran R}}$(Beijing) = 1.

A relation R is **completely** specified if dom R = X, otherwise it is *incompletely* specified. Consider X = {x_1,x_2,x_3} and Y = {y_1,y_2,y_3}. Let us also define R_1 = {(x_1,y_2), (x_2,y_1), (x_2,y_3)}; R_2 = {(x_1,y_1), (x_2,y_1), (x_3,y_1)}; and R_3 = {(x_2,y_3), (x_3,y_1)}. Consider the *three diagrams shown*. If each member of the domain appears **once** in R, then R is called a <u>mapping</u> or a <u>function</u>.



<u>Example</u>: Consider the two diagrams shown. Is R *injective, bijective and surjective*? In A, the answers are **no, no and no!** So R is incompletely specified. In B, it is (1) *completely* specified, (2) *injective* and (3) *surjective*. These three things also imply that it is **bijective**. (Note: *Surjective* = covers *all* y).



<u>Definition</u>: The *inverse* of a crisp relation R(X,Y) is a crisp relation $R^{-1}(X,Y)$, where $R^{-1}(X,Y)$ = {(y,x) | x$\in$X, y$\in$Y, (x,y)$\in$**R**}. Example: if R = {(x_1,y_1), (x_1,y_2), (x_3,y_2)}, then R^{-1} = {(y_1,x_1), (y_2,x_1), (y_2,x_3)}. In *matrix* form, $R = [^{1}0_{0}\ ^{1}0_{1}]$, and $R^{-1} = [^{1}{}_{1}\ ^{0}{}_{0}\ ^{0}{}_{1}]$. (The *columns* are the y_i while the *rows* are the x_i).

Notes: $R^{-1}(X,Y) = (R(X,Y))^t$. (RHS = the *matrix*). dom $R^{-1}(X,Y)$ = {y_1,y_2}. dom R(X,Y) = {x_1,x_3}. ran $R^{-1}(X,Y)$ = {x_1,x_3}. ran R(X,Y) = {y_1,y_2}. dom $R^{-1}(X,Y)$ = ran R(X,Y). dom R(X,Y) = ran $R^{-1}(X,Y)$. $(R^{-1})^{-1}$ = R.

$$\begin{bmatrix} 0.6 & 0.4 & 0.5 & 0 \\ 0.3 & 0 & 0 & 0 \\ 0.1 & 0.9 & 0.1 & 0.6 \\ 0.7 & 0 & 0.9 & 0 \end{bmatrix}$$

Example: X = {Housewives, Students, Middle aged business people, Eldery}; Y = {Yoghurt, Cars, Cat Food, Pension Schemes}; and R = "is interested in", a **fuzzy** relation. We could have a matrix as shown, where the *columns* correspond to Y, and the *rows* to X. For example, position (2,2) is "how much are students **interested** in cars?". $R^{-1}(X,Y)$ = "is interesting *for*".

Composition of relations. Consider that we have crisp sets X, Y and Z; and *binary* relations P(X,Y) and Q(Y,Z). The composition of P(X,Y) and Q(Y,Z) is a <u>relation</u> R(X,Z) = P(X,Y)$\circ$Q(Y,Z). (Y is *absorbed*). For crisp P and Q, R(X,Z) = {(x,z) | $\exists$ y$\in$Y such that (x,y)$\in$P and (y,z)$\in$Q}.

Properties of the composition: (1) *Not* commutative; (2) Distributive with **respect** to the inverse: $(P(X,Y)\circ Q(Y,Z))^{-1} = Q^{-1}(Y,Z)\circ P^{-1}(X,Y)$; (3) *Associative*: $P\circ(Q\circ R) = (P\circ Q)\circ R$). For fuzzy relations, the most *popular* definition is the max-min composition. Let P(X,Y) and Q(Y,Z) be **fuzzy** relations. The *max-min* composition, $\circ$, is a fuzzy relation R defined as $\mu_R(x,z) = \max_{y\in Y} \min\{\mu_P(x,y)\mu_Q(y,z)\}$.

Example: Z = Price; Z = {low, medium, high}; Q = "expensive". Now let P = "well off", P(X,Z). P = R$\circ$Q is "is interested in expensive". Looks like a matrix multiplication, as shown on the right.

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.1 & 0.9 \\ 1 & 0 & 0 \\ 0 & 0.6 & 0.4 \end{bmatrix};\ P = \begin{bmatrix} 0.6 & 0.4 & 0.5 & 0 \\ 0.3 & 0 & 0 & 0 \\ 0.1 & 0.9 & 0.1 & 0.6 \\ 0.7 & 0 & 0.9 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.1 & 0.9 \\ 1 & 0 & 0 \\ 0 & 0.6 & 0.4 \end{bmatrix}$$

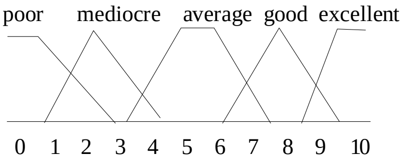
$$\begin{bmatrix} 0.6 & 0.1 & 0.4 \\ 0.3 & 0 & 0 \\ 0.1 & 0.6 & 0.9 \\ 0.9 & 0 & 0 \end{bmatrix}$$

We calculate P by finding the **maximum** of each “product”. For position (1,1), we have (0.6, 1.0), (0.4, 0), (0.5, 1) and (0,0). So we take the *maximum* of the minimums i.e. take 0.6.

Q: Carry out an **experiment**, and design a *membership function* for the set “student’s pocket” on the axis x = money. Explain the experiment and the membership function design. A: It is best that we use the *interval* method and ask each member of a group what they believe a student carries in their pocket. From the **interval** data, determine the **core** (*the intersection* of all intervals) and the **support** (the *union* of all intervals) of the fuzzy set.

Define the transition between *non membership* and *membership*. You can use a line segment or can take into account the interval **overlap**. For example, the function can be stepwise, such that for all x, $\mu_A(x) = N_Y/N$, where N_Y is the number of intervals *containing* x, and N is the *total* number of intervals (= the number of members of the group).

Q: Design a 10 question questionnaire to formulate a **numerical axis** for evaluating lecturers. Define on the new axis membership functions *corresponding* to “poor”, “mediocre”, “average”, “good” and “excellent”. A: All 10 questions should “**add**” in the same direction, e.g. Q1: “Is the lecturer *competent*”. Answers: No: 0, Yes, 1. The scale is from 0 to 10. We can also give weights, and make the *scale*, say, from 0 to 20.



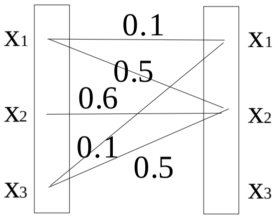
Q: The binary fuzzy relation R is defined on the sets X = {1,...,100} and Y = {51,...,100}, and represents the relation *“x is much smaller than y”*. The relation is defined by the following membership function: $\mu_R(x,y) = 1-{}^{x}/_{y}$ if $x \le y$, and $\mu_R(x,y) = 0$ otherwise, where $x \in X$ and $y \in Y$. dom R is a *fuzzy* set defined by $\mu_{\text{dom } R}(x) = 1-{}^{x}/_{100}$, for all $x \in \{1,2,...100\}$. ran R is a *fuzzy* set defined by $\mu_{\text{ranR}}(y) = 1-{}^{1}/_{y}$, for all $y \in \{51,...,100\}$. height R = $\max_x \max_y (1-{}^{x}/_{y}) = 1-{}^{1}/_{100} = 0.99$, so R is **subnormal**. As dom R ≠ l, R is *incompletely* specified.

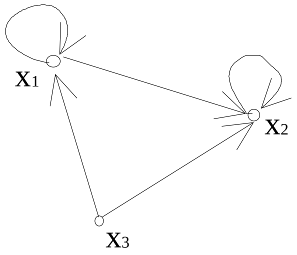
➢ 11th November 1999

Binary Relations (Crisp and Fuzzy) on a Single Set

In R(X,X), the ways to *express* R are by a Matrix, by a Saggital Diagrams, or by a Digraph. An example is shown below, where we have $X = \{x_1,x_2,x_3\}$:

$$\begin{bmatrix} 0.1 & 0.5 & 0 \\ 0 & 0.6 & 0 \\ 0.1 & 0.5 & 0 \end{bmatrix}$$

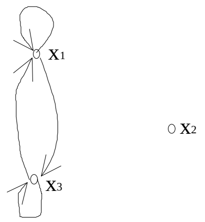




Properties of Binary Relations on a single set. ***Crisp***: Reflexivity: R(X,X) is reflexive iff $(x,x) \in R$ for all $x \in X$. If this doesn’t *hold*, then R is called **irreflexive**. If $(x,x) \neq R$, for all $x \in X$, then R is called *anti-reflexive*. So for $[{}^{1}0_{0}\ {}^{0}1_{0}\ {}^{1}0_{0}]$, it is irreflexive [$(x_1,x_1)$ yes; (x_2,x_2) yes; but (x_3,x_3) no]. $[{}^{0}1_{1}\ {}^{1}0_{1}\ {}^{0}1_{0}]$ is *anti-reflexive*.

2: **Symmetricity**. R is symmetric iff $\forall\ (x,y) \in R \Rightarrow (y,x) \in R$. Otherwise, R is called **asymmetric**. R is *anti-symmetric* iff $(x,y) \in R$ and $(y,x) \in R$ implies $x = y$. 3: **Transitivity**: R is *transitive* iff $(x,y) \in R$ and $(y,z) \in R \Rightarrow (x,z) \in R$. Otherwise, R is *non-transitive*. R is anti-transitive iff $(x,y) \in R$ and $(y,z) \in R \Rightarrow (x,z) \notin R$.

Example: $[^{1}0_{1}\ ^{0}0_{0}\ ^{1}0_{1}]$ is *irreflexive, symmetric and transitive*. Transitivity Check: $(x_1,x_3)(x_3,x_3) \Rightarrow (x_1,x_3)$ (*yes*); $(x_3,x_1)(x_1,x_1) \Rightarrow (x_3,x_1)$ (*yes*); $(x_3,x_1)(x_1,x_3) = (x_3,x_3)$ (*yes*); $(x_1,x_3)(x_3,x_1) \Rightarrow (x_1,x_1)$ (*yes*).



To specify a *relation* R (*to characteristic or describe it*), (1) Give **dom R**, **ran R**; (2) Is R *completely* specified or *incompletely* specified?; (3) Is R a function? — if yes, comment on **injective**, **surjective**, **bijective**; (4) *Reflexivity;* (5) *Symmetricity;* (6) *Transitivity*. Vague Example: R = "*is a friend of*". This is **Reflexive**, **Symmetric** and **Non-transitive**. Think about this.

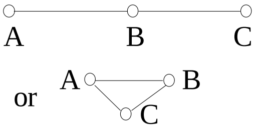
➢ 16th November 1999

Fuzzy Relations

Consider that we have R(X,X), with $\mu_R(x_1,x_2) \in [0,1]$, and $x_1,x_2 \in X$. A fuzzy relation R(X,X) is *reflexive* iff $\mu_R(x,x) = 1$ for all $x \in X$. Otherwise, it is **irreflexive**. If for all $x \in X$, $\mu_R(x,x) < 1$, then R(X,X) is called **antireflexive**. Examples: Reflexive: $[^{1}{}_{0.3}\ ^{0.1}{}_{1}]$; irreflexive $[^{1}{}_{0}\ ^{0.1}{}_{0.3}]$; antireflexive (irreflexive) $[^{0.6}{}_{0.7}\ ^{0.6}{}_{0.2}]$.

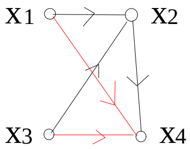
A relation R(X,X) is symmetric iff $\mu_A(x,y) = \mu_A(y,x)$ for all $x,y \in X$. Otherwise, R is called **asymmetric**. If (if $\mu_A(x,y) > 0$ and $\mu_A(y,x) > 0$ then $x = y$), then R is *anti symmetric*. Examples: symmetric: $[^{1}{}_{0.1}\ ^{0.1}{}_{0.3}]$. *What* is $[^{1}0_{0.1}\ ^{0.1}0_{0.2}\ ^{0.3}0_{0.6}]$? A *relation* R(X,X) is transitive (max-min transitive) iff $\mu_R(x,z) \geq \max_{y\in X} \min\{\mu_R(x,y), \mu_R(y,z)\}$; (*max-product* transitivity) iff $\mu_R(x,z) \geq \max_{y\in X} (\mu_R(x,y).\mu_R(y,z))$.

Let X be a set of *cities*, and R "very near". (1) **This is reflexive** (μ_R(Sofia, Sofia) = 1). (2) **This is symmetric** (μ_R(Sofia, Belgrade) = 0.8 = μ_R(Belgrade, Sofia)). (3) This is **not** transitive.



	Refl.	Anti Refl.	Symm.	Anti Symm.	Trans.
Crisp: Equivalence	Yes		Yes		Yes
Crisp: Quazi Equivalence			Yes		Yes
Crisp: (Tolerance) Compatibility. **Fuzzy**: Proximity	Yes		Yes		
Partial Ordering	Yes			Yes	Yes

The transitive **closure** of a crisp relation R(X,X) is the "*smallest*" relation on X.X which contains R. "smallest" = smallest cardinality. Example: $X = \{x_1,x_2,x_3,x_4\}$. $R(X,X) = \{(x_1,x_2), (x_3,x_2), (x_2,x_4)\}$. So $R_T(X,X) = \{(x_1,x_2), (x_3,x_2), (x_2,x_4), (x_1,x_4), (x_3,x_4)\}$. $|R_T(X,X)| = 5$.



An **algorithm** to find transitive closure: (1) Calculate $R' = R \cup (R \circ R)$. (2) If $R' \neq R$, assign $R = R'$, and do **step** one. Else (3) return R' as the *transitive* closure. So for the example on the previous page, following the algorithm, we <u>get</u>:

$$R \circ R = \begin{bmatrix} 0&1&0&0 \\ 0&0&0&1 \\ 0&1&0&0 \\ 0&0&0&0 \end{bmatrix} \begin{bmatrix} 0&1&0&0 \\ 0&0&0&1 \\ 0&1&0&0 \\ 0&0&0&0 \end{bmatrix} = \begin{bmatrix} 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \end{bmatrix}.\ R \cup (R \circ R) = \begin{bmatrix} 0&1&0&1 \\ 0&0&0&1 \\ 0&1&0&1 \\ 0&0&0&0 \end{bmatrix} = R'.$$

$$\text{Now } R \circ R = \begin{bmatrix} 0&1&0&1 \\ 0&0&0&1 \\ 0&1&0&1 \\ 0&0&0&0 \end{bmatrix} \begin{bmatrix} 0&1&0&1 \\ 0&0&0&1 \\ 0&1&0&1 \\ 0&0&0&0 \end{bmatrix} = \begin{bmatrix} 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \end{bmatrix}.\ R \cup (R \circ R) = \begin{bmatrix} 0&1&0&1 \\ 0&0&0&1 \\ 0&1&0&1 \\ 0&0&0&0 \end{bmatrix} = R. \text{ This is the same.}$$

<u>Assignment 2</u>

Q: Let A & B be **fuzzy** sets on U. Show by example that the *following* are possible: (try to find the "<u>minimal</u> possible" example, e.g. by keeping the cardinality of the universal set small): (a) $A \subset B$ and $\upsilon(A) = \upsilon(B)$; (b) $(\bar{A}') \neq (\bar{A})'$, where A' is the *closest crisp set* of A, and $\bar{A}$ is the *complement* of A; (c) $A \subset A^*$, where A* is a *sharpened* version of A.

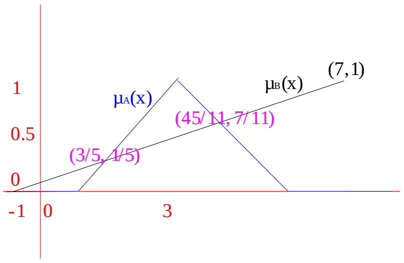
A: (a) Let $U = \{u_1\}$. *Define* A as $\mu_A(u_1) = 0.2$, and define B as $\mu_B(u_i) = 0.8$. So $A \subset B$ and $\nu(A) = \nu(B)$ holds. (b) Let $U = \{u_1\}$, and let $\mu_A(u_1) = 0.5$. Here, $\mu_{A'}(u_1) = 1$, and $\mu_{(\bar{A}')}(u_1) = 1$. Also, we have $\mu_{\bar{A}}(u_1) = 0.5$, and $\mu_{(\bar{A})'}(u_1) = 0$. So then $(\bar{A}') \neq (\bar{A})'$. (c) Let $U = \{u_1\}$, and let $\mu_A(u_1) = 0.6$. Define $\mu_{A^*}(u_1) = 0.8 > \mu_A(u_1)$. So $A \subset A^*$.

Q: Prove or **disprove** that (a) the bounded sum $A \oplus B$ is a t-conorm; (b) $\|A \oplus B\| \geq \|A \cup B\|$; (c) $S_1(A,B) = 0$ iff $S_2(A,B) = 0$. A: (a) $S(a,b) = \min\{1, a+b\}$. A1: **Limit** condition: $a,b \in [0,1]$, hence $a+b \geq 0$; $S(a,b) \geq 0$. Also, $S(a,b) \leq 1$ is obvious. A2: *Commutativity:* obvious. A3: Monotonicity: Let $a,b,c,d \in [0,1]$, with $a > c$ and $b > d$. Since $a+b > c+d$, then $\min\{1,a+b\} \geq \min\{1,c+d\}$.

A4: **Associativity**: Prove that $S(S(a,b), c) = S(a, S(b,c))$. LHS = $\min\{1, \min\{1,a+b\}+c\}$. RHS = $\min\{1, a+\min\{1,b+c\}\}$. Case 1: $a+b > 1$. Then LHS = 1. Case 1a: $b+c > 1$. Then RHS = 1. Case 1c: $b+c \leq 1$. Then RHS = $\min\{1, a+b+c\}$. **Because** $a+b > 1$, then RHS = 1. Case 2: $a+b \leq 1$. Then LHS = *min*$\{1,a+b+c\}$. Case 2a: $b+c > 1$. Then LHS = RHS = 1. **Case** 2b: $b+c \leq 1$. Then RHS = $\min\{1, a+b+c\}$. Since Cases 1 & 2 *contain all possibilities, we conclude that* LHS = RHS.

(b) Denote $a_i = \mu_A(u_i)$, and let $b_i = \mu_B(u_i)$. So $\|A \oplus B\| = {}^1\!/_n \Sigma_{i=1}^n \min\{1, a_i+b_i\}$; and $\|A \cup B\| = {}^1\!/_n \Sigma^{i=1}_n \max\{a_i, b_i\}$. For each pair of *corresponding terms,* the two summations $1 \geq a_i$ and $1 \geq b_i$ imply that $a \geq \max\{a_i, b_i\}$. Now $a_i+b_i \geq a_i$ and $a_i+b_i \geq b_i$ imply that $a_i+b_i \geq \max\{a_i,b_i\}$. Therefore, *for all* 1,...,n, $\min\{1,a_i+b_i\} \geq \max\{a_i, b_i\}$, and $\|A \otimes B\| \geq \|A \cup B\|$. (c) **Disprove**: Example: Let $U = \{u_1,u_2\}$; $\mu_A(u_1) = 0.3$; $\mu_A(u_2) = 0$; $\mu_B(u_1) = 0$; and $\mu_B(u_2) = 0.6$. Now $A \cap B = \phi$; $S_1(A,B) = 0$; and $S_2(A,B) = 1 - {}^1\!/_2(0.3+0.6) = 1 - 0.45 = 0.55$.

Q: Let A and B be *fuzzy* sets defined on **R** by the following membership functions: $\mu_A(x) = 1-{}^{1}/_{3}|x-3|$ if $x \in [0,6]$ (and 0 elsewhere); $\mu_B(x) = {}^{x+1}/_{8}$ if $x \in [-1,7]$ (and 0 elsewhere). Find the **implication** A⇒B in a functional form such as *above*, using (a) Gougen implication; (b) Bounded Product.



A: **Gougen** implication: 1 if $a \le b$, and ${}^{b}/_{a}$ if $a > b$. Find the **points** x_1 & x_2 for which $\mu_A(x) > \mu_B(x)$. So $1-{}^{1}/_{3}|x-3| = {}^{x+1}/_{8}$. (1): $x \in [0,3]$. Here, $1-{}^{1}/_{3}(3-x) = {}^{x+1}/_{8}$; $1-1+{}^{x}/_{3} = {}^{x+1}/_{8}$; $5x = 3$; $x_1 = {}^{3}/_{5}$. (2): $x \in [3,6]$. So $1-{}^{1}/_{3}(x-3) = {}^{x+1}/_{8}$; $1-{}^{x}/_{3}+1 = {}^{x+1}/_{8}$; $48-8x = 3x+3$; $11x = 45$; $x_2 = {}^{45}/_{11}$. The *implication* is a fuzzy set A ⇒ B with membership function $\mu_{A\Rightarrow B}(x) = {}^{3x+3}/_{24-8|x-3|}$ for $x \in ({}^{3}/_{5}, {}^{45}/_{11})$, and 1 *elsewhere*.

(b) **Bounded** product: max{0, a+b-1}. To find the *region*(s) where the bounded product is greater than zero, solve $a+b-1 > 0$. In our case, $1-{}^{1}/_{3}|x-3|+{}^{x+1}/_{8}-1 > 0$. (1): $x < 3$. We *have* $-{}^{3-x}/_{3}+{}^{x+1}/_{8} > 0$; $-24+8x+3x+3 > 0$; $11x > 21$; $x > {}^{21}/_{11}$. (2): $x > 3$. We have $-{}^{x-3}/_{3}+{}^{x+1}/_{8} > 0$; $-8x+24+3x+3 > 0$; $-5x > -27$; $x < {}^{27}/_{5}$. Therefore, $\mu_{A\Rightarrow B} = -{}^{|x-3|}/_{3}+{}^{x+1}/_{8}$ for $x \in ({}^{21}/_{11}, {}^{27}/_{5})$, and 0 *elsewhere*.

➢ 18th November 1999

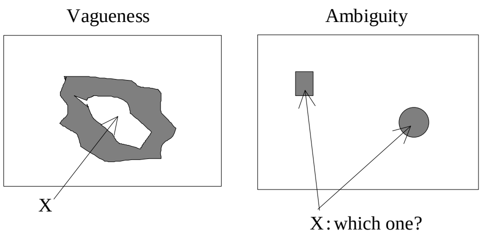
Transitive Closure of Fuzzy Relations

R' (the *transitive* closure) is obtained by the **same** algorithm as for crisp R. (1) $R\cup(R\circ R) = R$, then R is transitive (and is its *own* transitive closure). (2) $R' = R\cup(R\circ R)$; continue with step (1). *Example*: $R = [{}^{0}0.5_{0.6}\ {}^{0}0.2_{0}\ {}^{0.3}0.1_{0}]$. This is *anti-reflexive, asymmetric, and not transitive*. This is how we do the **transitive** part:

$R\circ R = [{}^{0}0.5_{0.6}\ {}^{0}0.2_{0}\ {}^{0.3}0.1_{0}]\circ[{}^{0}0.5_{0.6}\ {}^{0}0.2_{0}\ {}^{0.3}0.1_{0}] = [{}^{0.3}0.2_{0}\ {}^{0}0.2_{0}\ {}^{0}0.3_{0.3}]$ (*min-max composition*). $R\cup(R\circ R) = [{}^{0.3}0.5_{0.6}\ {}^{0}0.2_{0}\ {}^{0.3}0.3_{0.3}] = R'$. This is **not** transitive, but we can find the transitive *closure*: $R'\circ R' = [{}^{0.3}0.5_{0.6}\ {}^{0}0.2_{0}\ {}^{0.3}0.3_{0.3}]\circ[{}^{0.3}0.5_{0.6}\ {}^{0}0.2_{0}\ {}^{0.3}0.3_{0.3}] = [{}^{0.3}0.3_{0.3}\ {}^{0}0.2_{0}\ {}^{0.3}0.3_{0.3}]$. Now $R'\cup(R'\circ R') = [{}^{0.3}0.5_{0.6}\ {}^{0}0.2_{0}\ {}^{0.3}0.3_{0.3}] = R'$. So R' is the *transitive closure* of R.

Uncertainty and Information

Uncertainty has *many definitions in the dictionary*, but we are interested in **Vagueness** and **Ambiguity**. In the diagrams, *Vagueness* could be "X is sexy" (nonspecificity). Ambiguity could be that a witness says *"I saw somebody, but I'm not sure if it was a man or a woman"* (dissonance).



Two classical measures of uncertainty: **Hartley**, 1928, based on *set* theory; **Shannon**, 1948, based on *probability* theory.

Let $X = \{x_1, x_2, .., x_n\}$ be a set of *symbols*. A message of length S is a *sequence* of S symbols from X. The number of all possible messages of **length** s is n^s. Let $I(n^s)$ be the amount of information associated with S *selections* from X, where $|X| = n$. Now $I(n^s) = k(N).S$. [X_1, $|X_1| = n_1$; X_2, $|X_2| = n_2$]. When the *number* of sequences of length S_1 from X_1, and length S_2 from X_2, are such that $n_1^{S_1} = n_2^{S_2}$, the two sets should have the **same** information.

Now $k(n_1)S_1 = k(n_2).S_2$; $S_2/S_1 = k(n_1)/k(n_2)$; $S_1\log_b n_1 = S_2\log_b n_2$; $S_2/S_1 = \log_b n_1/\log_b n_2$; $k(n_1)/k(n_2) = \log_b n_1/\log_b n_2$; $k(n) = k_0\log_b n$ (where k_0 is a *constant*). $k_0 = 1$ and $b = 2$: BIT.

➢ 23rd November 1999

Recall that $I(n^S) = k_0 S\log_b n$. With $k_0 = 1$ and $b = 2$, we get $I(n^S) = S.\log_2 n$; $I(N) = \log_2 N$, "bit". (I(N) is the total *number* of messages). 1 bit of information is the **information** in N = 2 messages.

Hartley information can be characterised by the *following* axioms: (1) Additivity: Let M & N be integers ($M, N \in \mathbf{N}$). Then $I(N,M) = I(N)+I(M)$. (2) Monotonicity: $I(N) \leq I(N+1)$ for all $N \in \mathbf{N}$. (3) Normalisation: $I(2) = 1$. Theorem: The *function* $I(N) = \log_2 N$ is the **only** function which satisfies axioms 1-3.

Example: Let X and Y be *sets* of messages. As you can see from the table below, not all *combinations* are possible.

X \ Y	*Out of politics*	*Gave Birth to Triplets*	*Awarded an Oscar*	*Broke w/a g/friend*
William	No	No	No	Maybe
Dawn French	No	Maybe	Maybe	No
Lord Archer	Maybe	No	No	Maybe

3 Types of Hartley Information

(1) **Simple** information: $I(X) = \log_2|X|$ $(I(|X|))$; $I(Y) = \log_2|Y|$. (2) **Joint** information: A relation over X×Y specifies the connection between the *two sets of messages*. Example: $R = [{}^{0}0_{1}\ {}^{0}1_{0}\ {}^{0}1_{0}\ {}^{1}0_{1}]$. Here, $|R| = 5$; and $I(X,Y) = \log_2|R|$. (3) **Conditional** information: $I(X|Y) = \log_2{}^{|R|}/_{|Y|} = \log_2|R|-\log_2|Y|$; and $I(Y|X) = \log_2{}^{|R|}/_{|X|} = \log_2|R| - \log_2|X|$.

${}^{|R|}/_{|Y|}$ is the *average* number of elements of X that can be selected under the condition that another element of Y has already been selected. ${}^{|R|}/_{|Y|}$ in the example is ${}^{5}/_{4} = 1.25$. Notice also that $I(X|Y) = \log_2|R|-\log_2|Y| = I(X,Y)-I(Y)$; and $I(Y|X) = I(X,Y)-I(X)$.

Definition: The information transmission: $T(X,Y) = I(X)+I(Y)-I(X,Y)$. Definition: X & Y are called *non-interactive* if the selection of elements of X does not depend on Y (and vice-versa). (All combinations are possible). Let $R = X\times Y$. For *non-interactive* X & Y, $T(X,Y) = I(X)+I(Y)-I(X,Y) = \log_2|X|+\log_2|Y|-\log_2|R| = \log_2|X|+\log_2|Y|-\log_2|XY| = 0$.

For other X & Y, $T(X,Y) > 0$. For non-interactive X & Y, $T(X|Y) = I(X,Y)-I(Y) = I(X)+I(Y)-I(Y) = I(X)$; and $I(Y|X) = I(Y)$. Consider the **example** in the table below. There, $I(X) = \log_2 5 = 2.4$; $I(Y) = \log_2 3 = 1.6$; $I(X,Y) = \log_2 8 = 3$; $I(X|Y) = I(X,Y)-I(Y) = 3-1.6 = 1.4$; $I(Y|X) = 3-2.4 = 0.6$; $T(X,Y) = 2.4+1.6-3 = 1$; and $T(X,Y)_{max} = I(X+I(Y)$. ($|R| = 1$ so $\log_2 1 = 0$).

X / Y	*Appendicitis*	*Heart Disease*	*Pneumonia*
Fever (39°)	0	0	1
Sub febrile (37.1°)	1	0	1
Severe Pain	1	1	0
Cough	0	0	1
Smoking	0	1	1

Hartley information comes from set theory. (X) high means *many choices* - vagueness / non specificity.

Tutorial

Q: (a) Can a **reflexive** relation be **antitransitive**? Why? (Remember that antitransitivity is defined for *Crisp* relations only). (b) Can an **irreflexive** but not **antireflexive** relation be **antitransitive**? Why? (c) Prove that for any **transitive** relation R (crisp or fuzzy), $R\circ R \subseteq R$. A: (a) For *all* $x \in X$, $(x,x) \in R$ (reflexive). Further, $(x,y) \in R$ and $(y,z) \in R \Rightarrow (x,z) \notin R$ (*antitransitive*). Now form $(x,x) \in R$, and $(x,x) \in R$, $\Rightarrow (x,x) \notin R$. Contradiction! So a reflexive R cannot be antitransitive.

(b) If R is not antitransitive, then there is an $x \in X$ such that $(x,x) \in R$. Then the above argument applies; and R *cannot* be antitransitive. (c) The transitive closure algorithm implies that $R\cup Q = R$, where $R\circ R = Q$ (*define* it). So $\max\{\mu_R(x,y), \mu_Q(x,y)\} = \mu_R(x,y)$ for all $x,y \in X$. So $\mu_Q(x,y) \subseteq \mu_R(x,y)$. Therefore, $Q = R\circ R \subseteq R$.

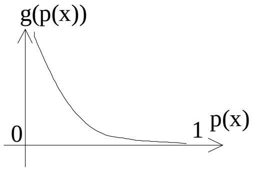
Q: Let $X = \{x_1, x_2, x_3\}$, and let R be a *fuzzy relation on X*, defined by $R = [{}^{0.1}0.2_{0.7}\ {}^{0.3}0.1_{0.2}\ {}^{0.4}0_{0.1}]$. (a) Find and *characterise* the 0.4 cut ($R_{0.4}$) of R. (b) Find $R^{-1}\circ R_{0.4}$. (c) Find a value of α such that R_α is a mapping (which could be *incompletely* specified). What are the values of **min** α and **inf** α?

A: (a) the 0.4 cut *matrix* is $[{}^{0}0_{1}\ {}^{0}0_{0}\ {}^{1}0_{0}]$. Therefore, dom $R_{0.4} = \{x_1, x_3\}$, and ran $R_{0.4} = \{x_1, x_3\}$. It is a *mapping; not* completely specified; injective; *not* surjective; *not* bijective; anti reflexive; symmetric; and (by algorithm) non-transitive. (b) Use the **transpose**. (c) α is in the range (0.3, 1]. It *is* a mapping with e.g. $\alpha = 0.4$. It has no minimum (it does not exist), but the *inf* is 0.3.

➢ 25th November 1999

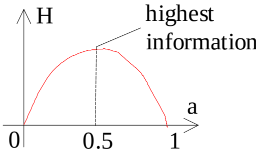
Shannon Entropy

Definition: $H(p(x) \mid x\in X) = -\Sigma_{x\in X}\, p(x).\log_2 p(x)$. (p(x) is the *probability mass function).* If $x \in X$ and $p(x) = 0.99$, there is **no** surprise if X occurs. If $p(x) = 0.01$, there is a *big* surprise if X occurs. Define g → "surprise" when X occurs. "surprise" = information; g: $[0,1]\rightarrow[0,\infty)$. Notes: (1) g is **monotone** decreasing; (2) g is **additive** w.r.t. the joint observations of independent events: $g(p(A,B)) = g(p(A).p(B))$ = (because A and B are independent) $= g(p(A))+g(p(B))$.



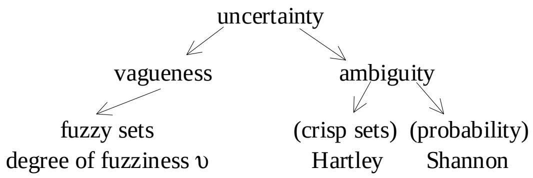
Cauchy equation: the solution is $g(t) = k\log_b t$ (k and b are constants). Now g(t) is monotone *decreasing* while $\log_b t$ is monotone *increasing*. Therefore, k should be **negative**. Pick $k = -1$ and $b = 2$, so that $g(t) = -\log_2 t$. Now g(p(x)) is the **information** if this x occurs. Let $X = \{x_1,\ldots,x_n\}$, with $x_i \in X$. What is the *expected* information of X?

The expected information of X is defined as $\Sigma_{x_i\in X}\, p(x_i)g(p(x_i)) = \Sigma_{x_i\in}$ $p(x_i)(-\log_2 p(x_i)) = -\Sigma_{x_i\in X}\, p(x^i)\log_2 p(x_i)$. Example: $X = \{x_1,x_2\}$; $p(x_1) = a$; $p(x_2$ = 1-a.



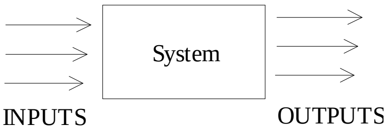
Properties: $0 \le H(p(x) \mid x \in X)$. *Exploration*: If $p(x) = 1$ for some x, then $1.\log_2 1 = 0$. If $p(x) = 0$, then we have $0.\log_2 0 = ?$ Now $^{\lim}{}_{(p(x_i)\to 0)}\ p(x_i)\log_2 p(x_i) = {}^{\lim}{}_{(p(x_i)\to 0)}\ \log_2 p(x_i)/(1/p(x_i))$ = (by l'Hospital) = $^{\lim}{}_{(p(x_i)\to 0)}(1/p(x_i)\ln 2)/(-1/p(x_i)^2)$ = $^{\lim}{}_{(p(x_i)\to 0)}\ -p(x_i)/\ln 2 = 0$. *Maximal* value: $p(x_i) = {}^1/_n$ (*equiprobable* events). $H(p(x) \mid x \in X) = -\Sigma_{x_i \in X}\ {}^1/_n \log_2 {}^1/_n = -\log_2 {}^1/_n$ (sum over *n terms*). X is **discrete** and **finite**.

Boltzmann Entropy



Definition: $B(q(x) \mid x \in [a,b]) = \int_a^b q(x)\log_2 q(x)dx$ Notes: q(x) is the *probability density function*; Boltzmann entropy is **not** an extension of Shannon entropy.

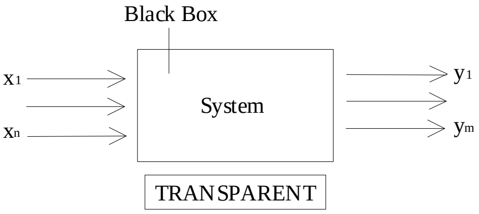
Fuzzy Systems



We have Inputs $x_1,\ldots,x_n$, and Outputs $y_1,\ldots,y_m$. If we have $n > 1$ and $m > 1$, then we have Multi-Input-Multi-Output (MIMO). If we have $n > 1$ and $m = 1$, then we have Multi-Input-Single-Output (MISO). If we have $n = 1$ and $m = 1$, then we have Single-Input-Single-Output (SISO).

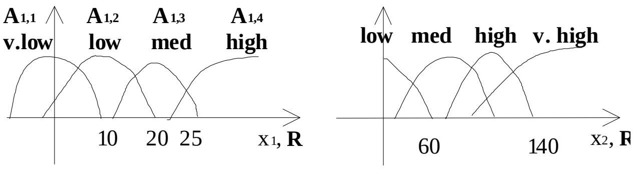
➢ 30th November 1999

Example: **Weather** Forecast. Consider that we have x_1, x_2 and x_3 *for "today"* parameters; x_4, x_5, x_6 and x_7 for "las week", and $x_8,\ldots,x_{20}$ for "the weather over Europe from the last week". Now $x_i \in \mathbf{R}$, $\underline{x} = [x_1,\ldots,x_n]^T \in \mathbf{R}^n$. Further, y_1 = tomorrow's min. temp., y_2 = tomorrow's max temp., $y_3 = \ldots$, $\underline{y} = [y_1,\ldots,y_m]^T \in \mathbf{R}^m$. And D, the system, is defined as D: $\mathbf{R}^n \to \mathbf{R}^m$.

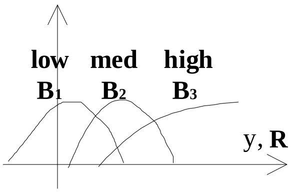


For all $\underline{x} \in \mathbf{R}^n$, the mapping D assigns to $\underline{x}$ an *output* vector $\underline{y} \in \mathbf{R}^m$. Fuzzy systems are usually based on IF-THEN rules. The "if" part connects the *inputs*; (the premise or the antecedent part); while the *"then"* part connects the outputs (the consequent part). Example: If x_1 is low, and x_2 is high, THEN y_1 is low, and y_2 is low. (x_1 = temperature; x_2 = speed of wind; y_1 = maximum temperature; and y_2 = minimum temperature tomorrow).

Approximate Reasoning. If x_1 is low and x is high. ("if": (clause 1, $\mu_{low}(x_1)$, $\mu_{low}(12) = 0.4$); "and": (clause 2, $\mu_{high}(x_2)$, $\mu_{high}(60) = 0.5$)). This is *conjunction*. Define $\tau(x) = \min\{\mu_{low}(x_1), \mu_{high}(x_2)\}$ (is the *firing* strength of the rule). Here, $\tau(x) = 0.4$, where $\underline{x} = [12, 60]^T$. The *consequent* can be of two types: (1) **Linguistic**; (2) **Functional:** $\underline{y} = f(\underline{x})$.



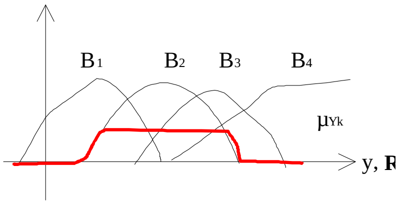
Put m = 1, and consider **MISO** and **SISO**. R_k: If x_1 is $A_{1,i(k,1)}$ and .... and x_n is $A_{n,i(k,n)}$, THEN y is $B_{0(k)}$ (*linguistic*). A and B are fuzzy sets corresponding to linguistic terms. Notation: In $A_{j,i(k,j)}$, j is the number of the **input**; i is th **indicator** function; and k is the number of the **rule**. Let, for our example the rule be R_5: If x_1 is $A_{1,i(5,1)}$ and x_2 is $A_{2,i(5,2)}$, then y is $B_{0(5)}$ (the indicato function for the *output*).

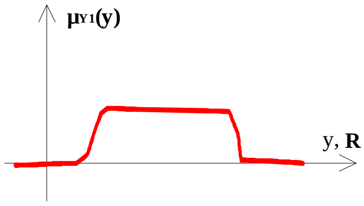


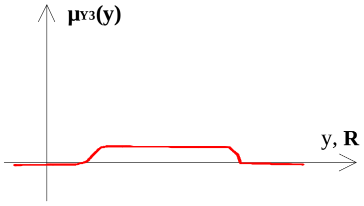
Mandani-Assilian

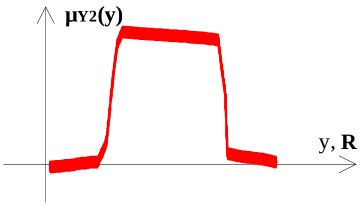
(1) Set up the *inputs, output and the rule base* ($x_1,...,x_n$; y; $R_1,...,R_L$). (2) Specify $\underline{x} = [x_1,...,x_n]^T$, the *concrete* input. (3) Find (from the **membership** functions)$\mu_{A_{j,i[k,j]}}(x_j)$. for all j = 1,...,n and k = 1,...,L. (4) Calculate the **firing** strength of each rule, with the following formula: $\tau_k(\underline{x}) = \min\{\mu_{A_{1,i(k,1)}}(x_1),...,\mu_{A_{n,i(k,n)}}(x_n)\}$ (k = 1,...,L). (5) *Implement* the implication. Y_k is the fuzzy output of the rule R_k. It is a fuzzy set defined on **R**. (Corresponding to the *output y).*

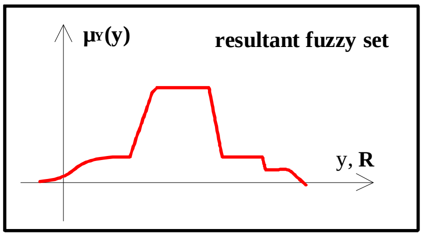
We "cut" the **consequent** set $B_{0(k)}$ with the firing strength $\tau_k(\underline{x})$. Therefore, $\mu_{Y_k}(y) = \min\{\tau_k(\underline{x}), \mu_{B_{0(k)}}(y)\}$, where $\tau_k(\underline{x})$ is a number in the interval [0,1]. Thus, we find $Y_1,...,Y_L$ (the *outpu* fuzzy sets). (6) **Aggregate** the output — use fuzzy sets $Y_1,...,Y_L$ to obtain the *resultant* fuzzy set Y.



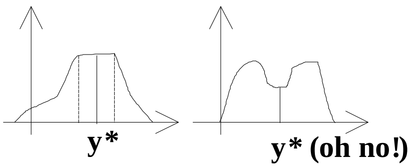








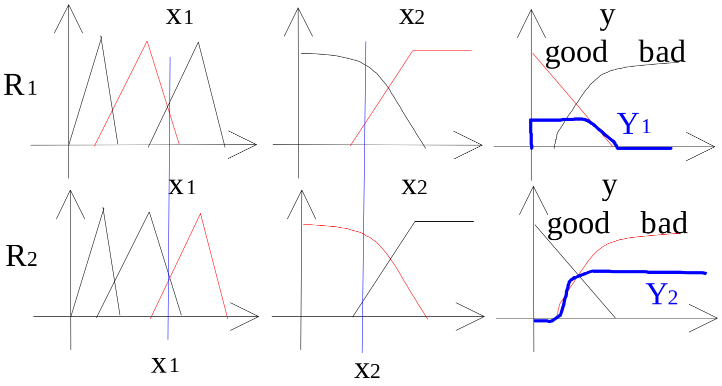
To **aggregate** the output (as shown on the left), use $\mu_Y(y) = \max\{\mu_{Y_1}(y),...,\mu_{Y_L}(y)\}$. (max = disjunction).

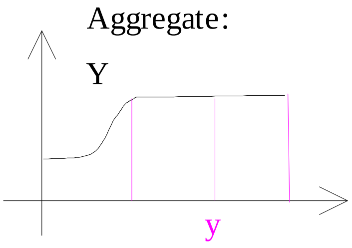


(7) **Defuzzify Y**. Use the *MOM* Method (The Mean of Maxima method, shown above); or the *COG* defuzzification. (Centre of gravity: $y \in \{V_1,...,V_S\}$ → universal set of output values. (Note: the output of the system is $y \in$ **R**)).

$$y^{**} = \frac{\sum_{i=1}^{S} \mu_Y(V_i).V_i}{\sum_{i=1}^{S} \mu_Y(V_i)} = \frac{\sum_{i=1}^{S} \mu_Y(V_i).V_i}{|Y|}$$

Example. (Inputs) x_1, x_2, y. (Outputs) $A_{1,1}$ — low, $A_{1,2}$ — medium, $A_{1,3}$ — high; $A_{2,1}$ — small, $A_{2,2}$ — large; B_1 — good, B_2 — bad. (Rule Base) R_1: If x_1 is medium and x_2 is large then y is good; R_2: If x_1 is high and x_2 is small then y is bad.





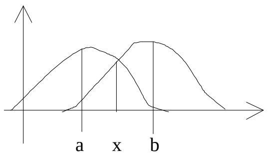
Assignment 3

Q: Let A & B be **fuzzy** sets on R, with $\mu_A(u) = \exp\{-{}^{1}/_{2}({}^{u-a}/_{\sigma})^2\}$, and $\mu_B(u) = \exp\{-{}^{1}/_{2}({}^{u-b}/_{\sigma})^2\}$. Knowing that *core*(A) = {1}; height(t(A,B)) = 0.88; (where t is the product t-norm); and $\mu_{i(A,B)}(0) = 0.37$, (where i(A,B) is the Gougen implication), find (a) a, b and σ; (b) C(A,B), where C is the consistency index; and (c) $\mu_{A\nabla B}(a)$, where A∇B is the symmetric difference based on *Hamming* distance.

A: (a) From core(A) = {1}, a = 1. From height(t(A,B)) = 0.88, $\mu_{t(A,B)}(u) = \exp\{-{}^{1}/_{2}({}^{u-1}/_{\sigma})^2-{}^{1}/_{2}({}^{u-b}/_{\sigma})^2)$ = (*simplifying*) = $\exp\{-{}^{1}/_{\sigma^2}(u-{}^{1+b}/_{2})^2+({}^{1-b}/_{2})^2)$. Since this is *positive* for any u, the **maximal** value of $\mu_{t(A,B)}$ is found for the **minimal** value of the thing in the exp{}. The minimal value of the contents of exp{} is therefore reached for $u = {}^{1+b}/_{2}$. From this, we *get* the formula $b = 1+2\sigma\sqrt{}-\ln(0.88)$.

Consider $\mu_{i(A,B)}(0) = 0.37$. In our problem, $\mu_B(0)/\mu_A(0) = 0.37$ from the *definition* of the implication. Therefore, $\exp\{-{}^{1}/_{2}(({}^{-b}/_{\sigma})^2-(-{}^{1}/_{\sigma})^2) = 0.37$; $-{}^{1}/_{2}.{}^{1}/_{\sigma^2}(b^2-1) = \ln 0.37$; ...; $b = \sqrt{}(1-2\sigma^2\ln 0.37)$. To find b and σ, solve the two equations *simultaneously* to give σ ≈ 0.9682 and b ≈ 1.6923.

(b) Consider height t(A,B), with t = min. (C(A,B) = height(A∩B)) Now x is the point of *intersection*, so equate $\mu_A(u)$ and $\mu_B(u)$ to give x = 1.34. Therefore, $C(A,B) = \exp\{-{}^{1}/_{2}({}^{1.34-1}/_{0.9682})^2\} \approx 0.94$. (c) $\mu_{A\nabla B}(1) = |\mu_A(1)-\mu_B(1)| = |\exp\{0\}-\exp\{-{}^{1}/_{2}({}^{1-1.6923}/_{0.9682})^2\}| = 1-0.07744 = 0.2256$.



Q: Let $A_1,...,A_n$ be fuzzy sets on U. Let us denote the pessimistic aggregation obtained by the *minimum* operation as A_{min}; and the optimistic aggregation obtained by the **maximum** as A_{max}. If A_{OWA} is obtained by an OWA aggregation operator, **prove** that $A_{min} \subseteq A_{OWA} \subseteq A_{max}$.

A: Let $u_j \in U$ be some *element* of U; and let $\mu_{A_i}(u_j) = a_i$. Then, for u_j, we have $\mu_{A_{min}}(u_j) = \min\{a_1,...,a_n\}$; $\mu_{A_{max}}(u_j) = \max\{a_1,...,a_n\}$; and $A_{OWA}(u_j) = \Sigma^n_{i=1} b_i a_{j(i)}$. (Note: this is for $b_i \geq 0$ and i = 1,...,n, where $\Sigma^n_{i=1} b_i = 1$, and $a_{j(1)},...,a_{j(n)}$ are the elements $a_1,...,a_n$, sorted in *descending* order).

Let $a_{min} = \mu_{A_{min}}(u_j) = \min\{a_1,...,a_n\}$; and let $a_{max} = \mu_{A_{max}}(u_j) = \max\{a_1,...,a_n\}$. ∀ set of OWA *coefficients* $b_1,...,b_n$, we have $\Sigma^n_{i=1} b_i a_{j(i)} \leq \Sigma^n_{i=1} b^j a_{max} = a_{max}$. Therefore, $\mu_{A_{OWA}}(u_j) \leq \mu_{A_{max}}(u_j)$ for all $u_j \in U$. *Similarly*, $\Sigma^n_{i=1} b_i a_{j(i)} \geq \Sigma^n_{i=1} b_i a_{min} = a_{min}$. Therefore, $\mu_{A_{min}}(u_j) \leq \mu_{A_{OWA}}(u_j)$.

Q: For which values of x is the following *fuzzy binary relation transitive*: $\begin{bmatrix} x & 0.2 \\ 0.1 & 0.6 \end{bmatrix}$. A: For transitivity to hold, we must have $R = R\cup(R\circ R)$. Now $R\circ R = \begin{bmatrix} x & 0.2 \\ 0.1 & 0.6 \end{bmatrix}\begin{bmatrix} x & 0.2 \\ 0.1 & 0.6 \end{bmatrix} = \begin{bmatrix} \max\{x,0.1\} & \max\{\min\{x,0.2\},0.2\} \\ \max\{\min\{x,0.1\},0.1\} & 0.6 \end{bmatrix}$. Consider max{min{x,a},a}, for a ∈ [0,1]. If x < a, the result is a. If x ≥ a, the result is *again* a.

Therefore, $R\circ R = \begin{bmatrix} \max\{x,0.1\} & 0.2 \\ 0.1 & 0.6 \end{bmatrix}$; $R\cup(R\circ R) = \begin{bmatrix} \max\{x,\max\{x,0.1\}\} & 0.2 \\ 0.1 & 0.6 \end{bmatrix}$. So max{x,max{x,0.1}} = max{x, 0.1}. For transitivity to *hold*, we must have max{x, 0.1} = x, i.e. x must satisfy x ≥ 0.1.

Q: For each of the following **binary** relations on a single set, state whether the relation is *reflexive, irreflexive or antireflexive*; *symmetric, asymmetric or antisymmetric*, and *transitive, nontransitive or antitransitive*. (i) "is a **sibling** of", (ii) "is a **parent** of", (iii) "is **smaller** than", (iv) "is the **same** height as", (v) "is as least as **tall** as", (vi) "**looks** like" (is similar but not necessarily *identical* to).

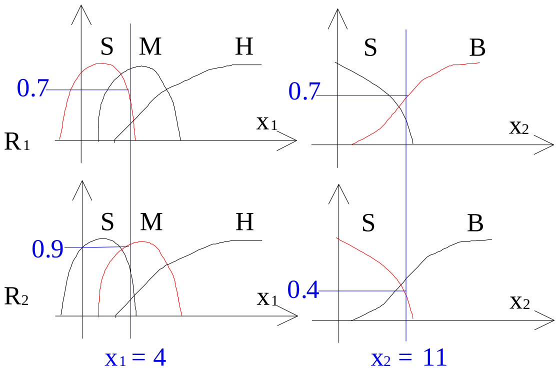
	R	IR	AR	S	AS	AnS	T	NT	AT
is a *sibling* of			✓	✓			✓		
is a *parent* of			✓			✓			✓
is *smaller* than			✓			✓	✓		
is the *same* height as	✓			✓			✓		
is as least as *tall* as	✓				✓		✓		
looks like (is similar but...)	✓			✓				✓	

➢ 2nd December 1999

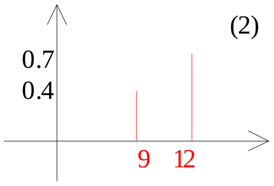
(**MA** *linguistic* fuzzy systems; **TSK**(TS) Takagi-Sugeno-Kang *functional* fuzzy systems). R_k: If x_1 is $A_{1,i(k,1)}$ and ... and x_n is $A_{n,i(k,n)}$, THEN $y_k = f_k(\underline{x})$. Usually, $y_k = b_{0k}+b_{1k}x_1+...+b_{nk}x_n$, where $b_{jk} \in$ **R** and j = 0,...,n (the *coefficients*).

(1) Build the system: Nominate $x_1,...,x_n$; Construct the *rule* base $R_1,...,R_L$; Pick $f_1,...,f_L$; and design the *membership* functions for the **inputs**. (2) For $\underline{x} \in \mathbf{R}^n$, find *all* $\mu_{A_{j,i(k,j)}}(x_j)$ for j = 1,...,n and k = 1,...,L. (3) Calculate the **firing** strengths $\tau_k(\underline{x}) = \min\{\mu_{A_{1,i(k,1)}}(x_1),...,\mu_{A_{n,i(k,n)}}(x_n)\}$. (min = *conjunction*) (4) Calculate the **outputs** $y_k = f_k(\underline{x})$ for k = 1,...,L. (5) **Fuse** the outputs using the formula: $y = \frac{\sum_{k=1}^{L} \tau_k(\boldsymbol{x}).y_k}{\sum_{k=1}^{L} \tau_k(\boldsymbol{x})}$.

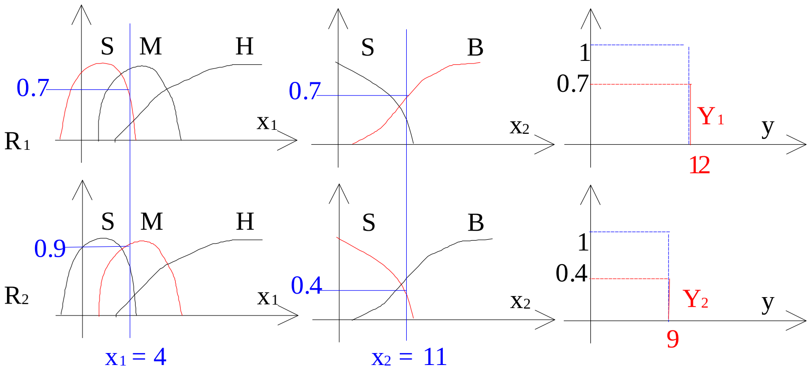
Example: $\underline{x} = [x_1\ x_2]^T$; R_1: If x_1 is small and x_2 is big, then $y_1 = 3x_1$; R_2: if x_1 is medium and x_2 is small then $y_2 = 5x_1-x_2$. From the diagrams, we see that $\tau_1(\underline{x}) = 0.7$; $y_1 = 3\times4 = 12$; $\tau_2(\underline{x}) = 0.4$; and $y_2 = 5\times4 - 11 = 9$. So x = $[4\ 11]^T$, and y = ${}^{0.7\times12+0.4\times9}/_{0.7+0.4}$ = 10.9.



The Equivalence Between the MA Model and the TSK Model



Assume that for $\underline{x} \in \mathbf{R}^n$ (input), all y_1 to y_L are different, so that $y_i \neq y_j$ for i ≠ j, with i, j ∈ {1,...,L}. (1) y_k define *singletons* on **R**. (2) Aggregate the $Y_1,...,Y_L$ by the *maximum method*. (3) *Defuzzify*: y = ${}^{9\times0.4+12\times0.7}/_{0.4+0.7}$.



➢ 9th December 1999

Revision

(1) **Fuzzy** sets, notions & notations. Elementary propositional logic. Membership functions. Measures of *fuzziness*.

(2) Operations on fuzzy sets. Inclusion, Similarities, 2 place operations (Set operations, ∪, ∩, t-norms, t-conorms). n-place operations *(aggregation, OWA)*. Implications.

(3) Fuzzy relations. Definitions, representation (**matrices**, **graphs**, **sets**). Relations on 2 sets and on 1 set. Mappings & properties. (**injective**, **surjective**, **bijective**). Characterisation of a relation. Properties of a binary relation on one set: *reflexivity, symmetricity, transitivity*.

(4) *Uncertainty* & *Information*. Hartley information, conditional information, information transmission. Shannon entropy.

(5) Fuzzy **IF-THEN** systems. Algorithms of operation. *Equivalence* between the two models.

<u>Questions</u>. If A is a fuzzy set defined over the *Cartesian* product of X = {1,2,3} and Y = {a,b}, then U = {(1,a), (1,b), (2,a), ..., (3,b)}. Given another set, we could then calculate $S_2 = 1-\|A\nabla B\|$. With OWA operations, consider that we apply OWA to 3 sets A, B & C, with coefficients (0.7, 0.3, 0). Then for every position, **<u>order</u>** the numbers, and then multiply by the coefficients. In the table shown, 0.24 = (0.3)(0.7)+(0.1)(0.3)+(0)(0); and 0.72 = (0.9)(0.7)+(0.3)(0.3)+(0.1)(0).

A	a	b	B	a	b	C	a	b	OWA	a	b
1	0.1	0.9	1	0	0.3	1	0.3	0.1	1	0.24	0.72
2	1	0.3	2	0.9	0.6	2	0	0	2		
3	0.2	0.6	3	0.4	0.1	3	1	0.3	3		

➢ 14th December 1999

Revision — Fuzzy Relations

Let X = {a,b}, and let R(X,X,X) be a *relation*, (a subset on the Cartesian product X×X×X), with U = X×X×X, and |U| = 8. This is a **ternary** relation, for example R = {(aab), (aba), (bbb)}. Crisp (R): 0 or 1; Fuzzy ($R^{\sim}$): Anything between 0 and 1. **Binary** Relations: R(X,Y) can be represented by *a matrix, a set, or a saggital diagram.*

<u>Example</u>: (Crisp): X = {a,b}; Y = {A,B,C}; and R = {(a,B), (b,A), (b,B)}. Here, dom R = X; ran R = {A,B}; and R is **not** a mapping, but **is** completely specified. Fuzzy example: R = $[{}^{0.3}{}_{0}\ {}^{0}{}_{0.6}\ {}^{1}{}_{0}]$, with R = {((a,A), 0.3), ((a,C), 1), ((b,B), 0.6)}. Here, dom R = {(a,1), (b,0.6)}; R is **not** completely specified; ran R = {(A,0.3), (B,0.6), (C,1)}; and R is **not** a mapping.

Relations on a single set X can also be shown with *digraphs*, e.g. X = {a,b} with R = {((a,a),0.6), ((b,a),0.7)}. Three properties: *Reflexivity, Symmetricity, Transitivity*. Note: with antisymmetricity, on each pair of off-diagonal terms, if one is non zero, then the other must be zero. For example, the following is reflexive & antisymmetric: $[{}^{1}0.3_{0}\ {}^{0}1_{0.4}\ {}^{0.1}0_{1}]$.

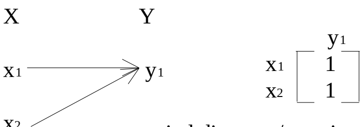
Fact: Something is **not** bijective if |X| ≠ |Y|. Q: If R is a *binary bijective* mapping, prove that X and Y are always interactive, i.e. prove that I(X,Y) ≠ I(X) + I(Y). A: We need |X| = |Y| = N, and |R| = N. Therefore, the above holds.

Assignment 4

Q: What is the meaning of *“completely specified”*, *“injective”*, *“bijective”*, *“surjective”* and *“mapping”*? Give an example where R is a **completely** specified, **surjective,** but non **bijective** mapping. Give an example where R **is not** a mapping, but is completely specified. A: A relation is *completely specified* if dom R = X; it is *injective* if the mapping is one-to-one only; it is *bijective* if all of Y is covered; it is *surjective* if it is completely specified, injective and bijective; and it is not a mapping if it is *one-to-many*.

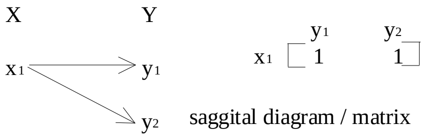
Example 1: Let X = $\{x_1,x_2\}$, and let Y = $\{y_1\}$. Therefore, X×Y = $\{(y_1,x_1), (x_2,y_1)\}$. Let us define a relation by R(X,Y) = $\{(x_1,y_1), (x_1,y_1)\}$. The relation is *completely specified* because domR = X; surjective because all of Y is covered; but not bijective because it is a *many-to-one* mapping. Example 2: Let X = $\{x_1\}$, and let Y = $\{y_1,y_2\}$. Let us define a relation by R(X,Y) = $\{(x_1,y_1), (x_1,y_2)\}$. The relation is completely specified because domR = X; but it is **not** a mapping because x_1 has links to y_1 & y_2, i.e. it is a *one-to-many* relation.

Example 1



saggital diagram / matrix

Example 2

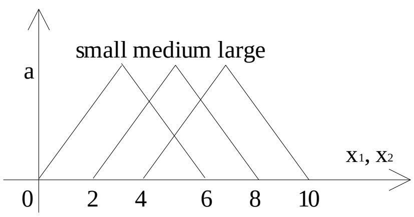


saggital diagram / matrix

Q: Prove that the *Shannon Entropy* H for U = $\{u_1,u_2\}$, where $p(u_1) = a$ and $p(u_2) = 1-a$, has a *maximum* for a = 0.5. Find this maximum. A: The *Shannon* entropy is defined as $H(p(u) \mid u \in U) = -\Sigma_{u\in U}\ p(u).\log_2 p(u) = -p(u_1).\log_2 p(u_1)+p(u_2).\log_2 p(u_2)) = -(a\log_2(a)+(1-a)\log_2(1-a)) = -a\log_2(a)-\log_2(1-a)+a\log_2(1-a)$. To find the *maximum points,* find the **stationary** points for $H(p(u) \mid u\in U)$.

Let $y = H(p(u) \mid u\in U)$, so that $y = -a\log_2(a)-\log_2(1-a)+a\log_2(1-a)$. Therefore, ${}^{dy}/_{da} = -\log_2(a)-1+{}^{1}/_{1-a}+\log_2(1-a)-{}^{a}/_{1-a}$. We find that this has a *stationary point* at a = ½, and we test ${}^{d^2y}/_{da^2}$ to find that this is a *maximum*. Now plug a = ${}^{1}/_{2}$ back into y to find out that this **maximum** value is $\log_2(2) = 1$.

Q: Find a relation between **two** sets of messages X & Y such that the information transmission is T(X,Y) = 3; X & Y are of the *same* cardinality; and the relation is *not* a mapping. A: $T(X,Y) = I(X)+I(Y)-I(X,Y) = \log_2|X|+\log_2|Y|-\log_2|R|$. Because X & Y are of the *same* cardinality, we must have I(X) = I(Y), and therefore we have $2\log_2|X| - \log_2|R| = \log_2({}^{|X|^2}/_{|R|})$. Now we **want** T(X,Y) = 3, so $3 = \log_2({}^{|X|^2}/_{|R|})$; $2^3 = {}^{|X|^2}/_{|R|}$, $8|R| = |X|^2$. We need to *choose* integers X and R satisfying the above. Choose |R| = 2, so |X| = 4. Choosing an **example** that is not a mapping, we can have the following: X = $\{x_1,x_2,x_3,x_4\}$; Y = $\{y_1,y_2,y_3,y_4\}$; and R = $\{(x_1,y_1), (x_1,y_2)\}$.



Q: A **MISO** fuzzy system is defined by the following rules: R_1: If x_1 is small and x_2 is small, then $y = 3x_1-x_2+5$. R_2: If x_1 is medium and x_2 is large, then $y = x_2-5$. x_1 & x_2 vary in [0,10], and membership functions for *small*, *medium and large* are as shown on the left. Express the membership functions in **functional** form, and calculate the output of the system for **x** = [5,8].

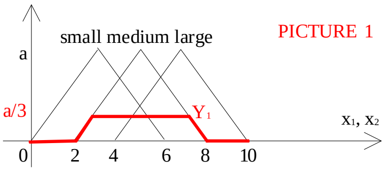
A: Denote the **peaks** of the *triangles* as having height a, where $a \in [0,1]$. Here are the *membership functions* in functional form: $\mu_{small}(x_1) = a({}^{1}/_{3}x)$ for $x_1 \in [0,3]$; $\mu_{small}(x_1) = a(-{}^{1}/_{3}x+2)$ for $x_1 \in [3,6]$; and $\mu_{small}(x_1) = 0$ for $x_1 \in [6,10]$. Now $\mu_{medium}(x_1) = 0$ for $x_1 \in [0,2]$; $\mu_{medium}(x_1) = a({}^{1}/_{2}x-{}^{2}/_{3})$ for $x_1 \in [2,5]$; $\mu_{medium}(x_1) = a(-{}^{1}/_{3}x+{}^{8}/_{3})$ for $x_1 \in [5,8]$; and $\mu_{medium}(x_1) = 0$ for $x_1 \in [9,10]$. Finally, $\mu_{large}(x_1) = 0$ for $x_1 \in [0,4]$; $\mu_{large}(x_1) = a({}^{1}/_{3}x-{}^{4}/_{3})$ for $x_1 \in [4,7]$; and $\mu_{large}(x_1) = a(-{}^{1}/_{3}x+{}^{10}/_{3})$ for $x_1 \in [7,10]$. The membership *functions* are the same for x_2.

We now have a TSK **functional** fuzzy system. Let us calculate the *output* of the system for $\underline{x} = [5,8]$. To start with, let us calculate the **firing** strength for each rule. Now $\tau_1(x) = \min\{\mu_{small}(x_1), \mu_{small}(x_2)\}$, so $\tau_1(5,8) = \min\{\mu_{small}(5), \mu_{small}(8)\} = \min\{a(-{}^{5}/_{3}+2), 0\} = 0$. And $\tau_2(x) = \min\{\mu_{medium}(x_1), \mu_{large}(x_2)\}$, so $\tau_2(5,8) = \min\{\mu_{medium}(5), \mu_{large}(8)\} = \min\{a({}^{5}/_{3}-{}^{2}/_{3}), a(-{}^{8}/_{3}+{}^{10}/_{3})\} = \min\{a(1), a({}^{2}/_{3}) = {}^{2a}/_{3}$.

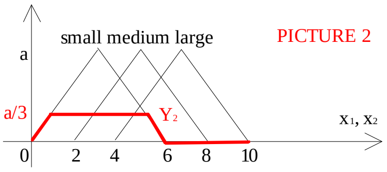
Let us now calculate the *outputs* $y_k = f_k(\underline{x})$ for $k = 1,\ldots,L$. To start with, $y_1 = 3x_1-x_2+5 = 3(5)-8+5 = 12$. Secondly, $y_2 = x_2-5 = 8-5 = 3$. To get the *output* for $\underline{x} = [5,8]$, let us fuse the *above* outputs, giving $y = {}^{0\times12+(2a/3)\times3}/_{0+(2a/3)} = {}^{2a}/_{(2a/3)} = 2a({}^{3}/_{2a}) = 3$.

Q: Consider an **MA** model with antecedents as above and *consequences*: R_1: .... then y is medium; R_2: ... then y is small. Assuming that y varies in [0,10], and that the *membership* functions are defined as in the previous question, calculate the output for **x** = [5,5] with MOM (mean of maxima) *defuzzification*.

A: Using the *Mandani-Assilian* method, we will now **calculate** the output for $\underline{x} = [5,5]$. To start with, let us calculate the *firing* strength for each rule. To start with, $\tau_1(\underline{x}) = \min\{\mu_{small}(x_1), \mu_{small}(x_2)\}$, so $\tau_1(5,5) = \min\{\mu_{small}(5), \mu_{small}(5)\} = \min\{a(-{}^{5}/_{3}+2), a(-{}^{5}/_{3}+2)\} = a(-{}^{5}/_{3}+2) = {}^{a}/_{3}$. Secondly, $\tau_2(\underline{x}) = \min\{\mu_{medium}(x_1), \mu_{large}(x_2)\}$, so $\tau_2(5,5) = \min\{\mu_{medium}(5), \mu_{large}(5)\} = \min\{a({}^{5}/_{3}-{}^{2}/_{3}), a({}^{5}/_{3}-{}^{4}/_{3})\} = a({}^{5}/_{3}-{}^{4}/_{3}) = {}^{a}/_{3}$. We shall now *implement* the implication, where Y_k is the fuzzy output of rule R_k.

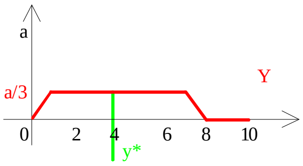


PICTURE 1



PICTURE 2

Here, $\mu_{Y_1}(y) = \min\{\tau_1(\underline{x}), \mu_{medium}(y)\}$. For $\underline{x} = [5,5]$, we have $\mu_{Y_1}(y) = \min\{\tau_1(5,5), \mu_{medium}(y)\} = \min\{{}^{a}/_{3}, \mu_{medium}(y)\}$. So we "cut" the *consequent* set "medium" with ${}^{a}/_{3}$ (as shown in picture 1). Now $\mu_{Y_2}(y) = \min\{\tau_2(\underline{x}), \mu_{small}(y)\}$. For $\underline{x} = [5,5]$, we have $\mu_{Y_2}(y) = \min\{\tau_2(5,5), \mu_{small}(y)\} = \min\{{}^{a}/_{3}, \mu_{small}(y)\}$. So we "cut" the *consequent* set "small" with ${}^{a}/_{3}$ (as shown in picture 2).



To finish, we *aggregate* the fuzzy sets Y_1 and Y_2 to obtain the *resultant* fuzzy set Y: $\mu_Y(y) = \max\{\mu_{Y_1}(y), \mu_{Y_2}(y)\}$ (as shown on the *left*). To defuzzify Y, we shall use the MOM method: a maxima strip occurs in [1,7], so y* occurs at ${}^{(1+7)}/_{2} = 4$.

Exam Paper: January 2000

SECTION 1 (Compulsory)

(1) (a) Define t-norm between two fuzzy sets giving a set of axioms. Give the name and a mathematical expression for each axiom. **[5 marks]**

(b) Let A be a fuzzy set on some universal set U. Let A' be the closest crisp set of A, $\bar{A}$ be the complement of A, A* be a sharpened version of A, and A_α be the α-cut of A. There are exactly four correct statements in the below table. Find and rewrite them (one mark each). (Make sure you write down *only four*.) **[4 marks]**

$A \subseteq A^*$	$\mathrm{core}(A) \subseteq A_{0.5}$	$A \subseteq A_{0.0}$
$A \subseteq A_{1.0}$	$A \subseteq \bar{A}$	$A \subseteq A'$
$A_{0.5} \subseteq A_{0.2}$	$A^* \subseteq A_{0.0}$	$\mathrm{core}(A) \subset A_{1.0}$

(c) The table below shows three fuzzy sets, A, B and C on the same universal set U.

U →	u_1	u_2	u_3	u_4	u_5
μ_A	0.1	0.4	0.9	0.3	0.2
μ_B	0.1	0.6	0.1	0.0	0.7
μ_C	0.1	0.0	0.0	0.2	0.4

(i) Aggregate the three fuzzy sets using OWA (ordered weighted averaging) coefficients [0, ½, ½]. **[2 marks]**

(ii) Find $|\mathrm{core}(A \cup B)|$ and $\mathrm{height}(A \cup \bar{A})$. **[1 mark]**

(iii) Calculate the fuzziness $\nu(A)$ of A. **[1 mark]**

(iv) Calculate the similarity between A and B using a measure of similarity of your choice. **[2 marks]**

(d) (i) Draw membership functions for the linguistic labels “small” and “large” for $x_1 \in [5, 50]$. Draw membership functions for the linguistic labels “early” and “approximately on time” on a time scale $x_2 \in [-10, 10]$. **[2 marks]**

(ii) Formulate all possible if-then rules for the two variables and calculate their firing strengths for input $[20, -4]^T$ (you can estimate approximately the degree of membership from your graph). **[3 marks]**

SECTION 2 (Answer 2 out of 4 questions)

(2) Prove that the Hamacher product $T(a, b) = \frac{ab}{a+b-ab}$ is a t-norm. **[15 marks]**

(3) Let R(X, X) be a relation on the (crisp) set X = {x_1, ..., x_4} expressed as the matrix

$$R = \begin{bmatrix} 0.1 & 0.6 & 0.4 & 0.1 \\ 0.9 & 0.6 & 0.1 & 0.3 \\ 1.0 & 0.0 & 0.0 & 0.4 \\ 0.0 & 0.1 & 0.5 & 0.8 \end{bmatrix}$$

(a) Find an α such that the α-cut of R has the maximal possible cardinality and is a *mapping* (could be incompletely specified). Characterise the mapping obtained. **[3 marks]**

(b) Characterise the relation $R_{0.6}$ (the 0.6-cut of R). **[3 marks]**

(c) Prove that antitransitivity involves antireflexivity (for any relation). **[5 marks]**

(d) Calculate the max-min composition $R_{0.3} \circ R^{-1}$, where R^{-1} is the inverse of relation R. **[4 marks]**

(4) Let X and Y be sets of messages.

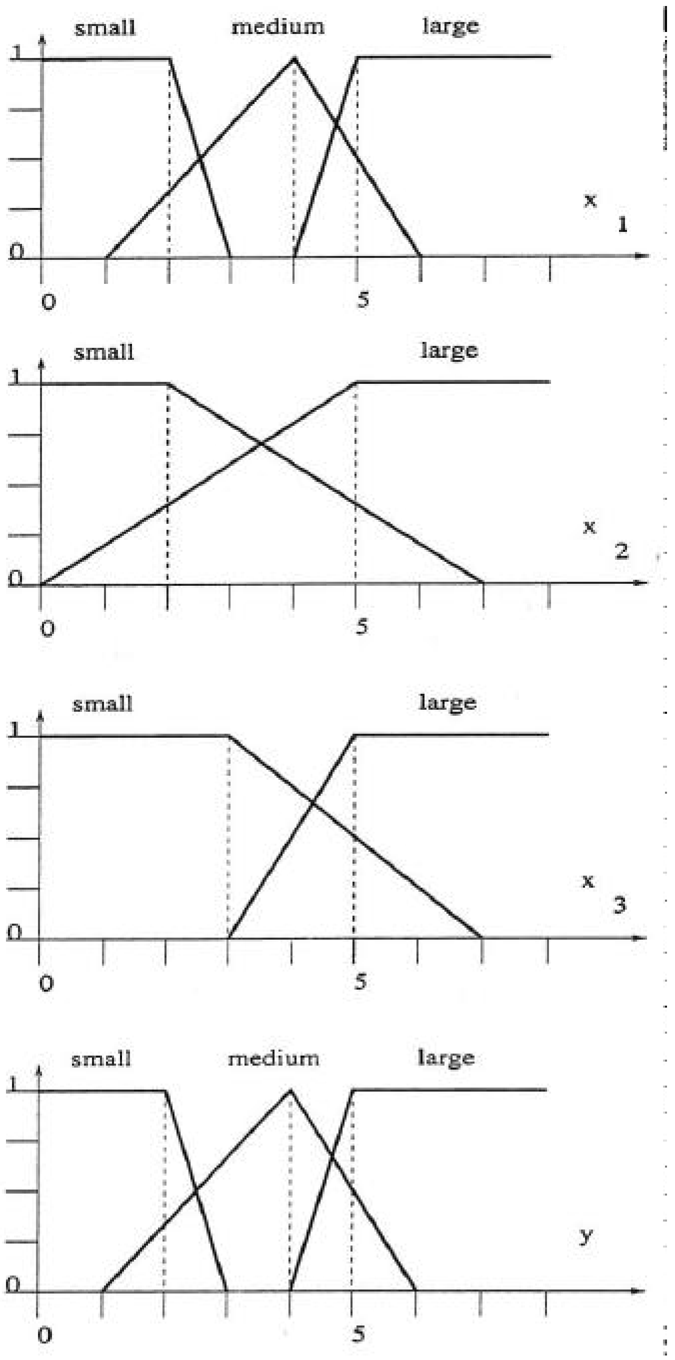
(a) Let R(X, Y) be a crisp binary relation which defines a completely specified bijective mapping f: X→Y. Knowing that the cardinality of X is N, find the information transmission T(X, Y). **[3 marks]**

(b) Give an example of a crisp relation R(X, Y) that defines a completely specified injective mapping f: X→Y with joint information value I(X, Y) = 2. **[4 marks]**

(c) Let |X| = N and |Y| = M. What are the minimal and the maximal possible values of the joint information I(X, Y) if R defines a completely specified surjective mapping f: X→Y? **[4 marks]**

(d) Let x be a random variable taking values in the set X = {-1, 0, 1, 2, 3, 4}. Knowing that the expected value (the mean) of x is 3, suggest a probability mass function with minimal value of the Shannon entropy (no proof is required). **[4 marks]**

(5) (a) Assume you have a MISO MA-type fuzzy system. The membership functions partitioning the input feature axes as given in the figure below.



(i) *Calculate* the firing strength of the rule:
IF x_1 is *medium* AND x_2 is *large* AND x_3 is *small* THEN y is *medium* for input $[4, 2, 1]^T$ and product as the AND operation. **[4 marks]**

(ii) Draw the output fuzzy set for this rule using minimum as the implication. **[3 marks]**

(b) The firing strength of the 4 rules of a SISO TSK fuzzy system for input x = 6 are respectively 0.4, 0.7, 0.1 and 0.2. The consequent parts are as follows:

R_1: ... THEN $y = 3x$
R_2: ... THEN $y = 2x-1$
R_3: ... THEN $y = x+1$
R_4: ... THEN $y = 3x$

(i) Find the output of the system. **[3 marks]**

(ii) Assume that the consequents are normal singletons at the respective values of y, and the system is an MA model. Find the output fuzzy set of the system and calculate the output by the Mean Of Maximums (MOM) defuzzification method. **[5 marks]**

(Questions done: 1, 2, 5)