FUZZY SYSTEMS

Jaya Sil
Indian Institute of Engineering Science and Technology, Shibpur
Howrah-711103

Uncertainty

- The precision of mathematics owes its success due to the efforts of Aristotle and his predecessors.

- Their efforts led to a concise theory of logic and Mathematics.

- The “Law of the Excluded Middle,” states that every proposition must either be True or False.

- There were strong and immediate objections, Heraclitus proposed that things could be simultaneously **True and not True.**

Uncertainty (19th Century)

•The scientific view at the time is summarized by Scottish physicist and mathematician William Thomson as “when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind".

•In 1871, Austrian physicist L. Boltzmann introduced statistical thermodynamics, a theory developed to study the motions of gas molecules in a closed space.

•Followed by the establishment of a new field of physics which applied the same statistical reasoning in a mechanical context: statistical mechanics.

Uncertainty (20th Century)

- The main difference between these new and traditional science disciplines was the use of statistics (and probability theory) in order to avoid the problems of Newtonian mechanics which were not suitable for the complex systems at hand.
- As Klir notes “When statistical mechanics was accepted, by the scientific community as a legitimate area of science at the beginning of the 20th century, the negative attitude toward uncertainty was for the first time revised. Uncertainty became recognized as useful, or even essential, in certain scientific enquiries.”

Uncertainty

- While uncertainty became firmly established as a part of science, it was taken for granted that all occurrences of uncertainty could be handled by **probability theory.**

- During the second half of the 20th century, it quickly became apparent that fundamentally new mathematical models and methods were required as the classical - no uncertainty or uncertainty modelled through probability could not address the uncertainty as it was present in the huge class of real world problems.

History

- Plato laid a foundation for fuzzy logic, indicating that there was a third region (beyond True and False).
- Lukasiewicz proposed a systematic alternative to the bi-valued logic of Aristotle describing a three-valued logic where the third value translated as “possible,” and assigned a numeric value between True and False.
- Later, he explored four-valued, five-valued logics, and declared that in principle there was nothing to prevent the derivation of an infinite-valued logic.
- Knuth proposed three-valued logic,the integral range [-1,0,+1]

Fuzzy Logic

- Fuzzy logic was developed by Lotfi A. Zadeh in the 1960s in order to provide mathematical rules and functions which permitted natural language queries and uncertainty handling.
- Fuzzy Logic attempts to mimic the way of human thinking to reason in an approximate way rather than a precise way.
  - For example, we don't say "If the temperature is above 24 degrees and the cloud cover is less than 10% and I have 3 hours time, I will go for a hike with a probability of 0.47."
  - We say "If the weather is nice and I have a little time, I will probably go for a walk."
  - Fuzzy logic attempts to provide a way of mathematically expressing the uncertainty of information.
- Fuzzy Logic Control can be regarded as a way of converting linguistic control information to mathematical control information.

Fuzzy Logic

- Zadeh developed fuzzy logic as a way of processing data.
- Instead a data element to be either a member or non-member of a set, he introduced the idea of partial set membership.
- In 1974 Mamdani and Assilian used fuzzy logic to regulate a steam engine.
- In 1985 researchers at Bell laboratories developed the first fuzzy logic chip.

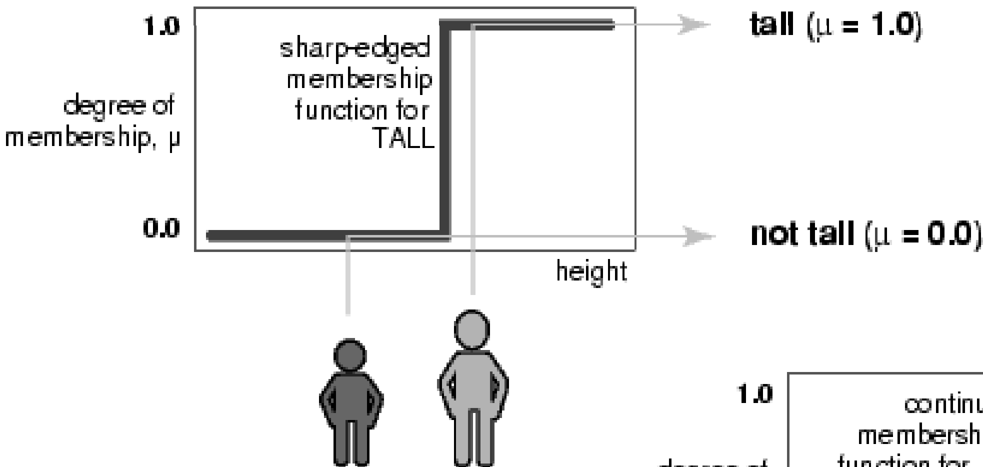
Motivation

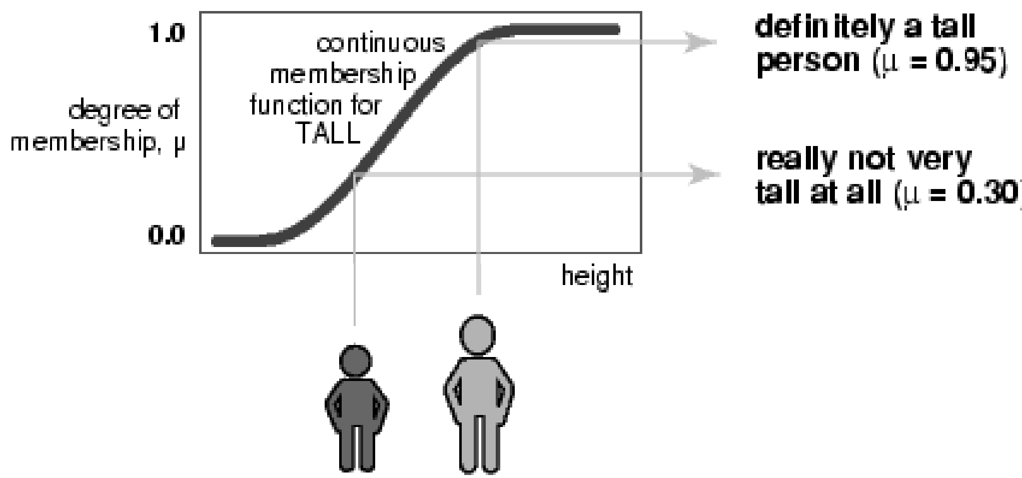
- Boolean logic assumes that every fact is either entirely true or false.
- The term “fuzzy logic” refers to a logic of approximation.
- Fuzzy logic allows for varying degrees of truth.
- Computers can apply this logic to represent vague and imprecise ideas, such as “hot”, “tall” or “good”.

Introduction

- Natural language employs many vague and imprecise concepts.
- It might be argued that vagueness is an obstacle to clarity of meaning.
- Translating such statements into more precise language removes some of their semantic value.
- This is what happens when natural language is translated into classic logic.
- The loss is severe when the issue is related to queries and knowledge.

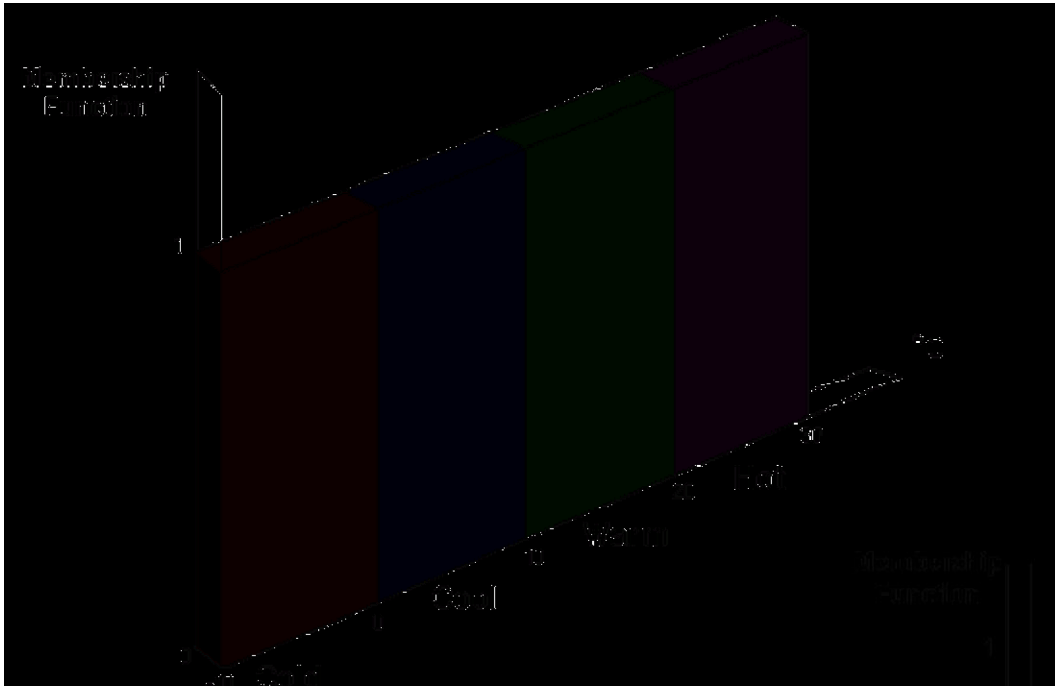
What Is Lost……

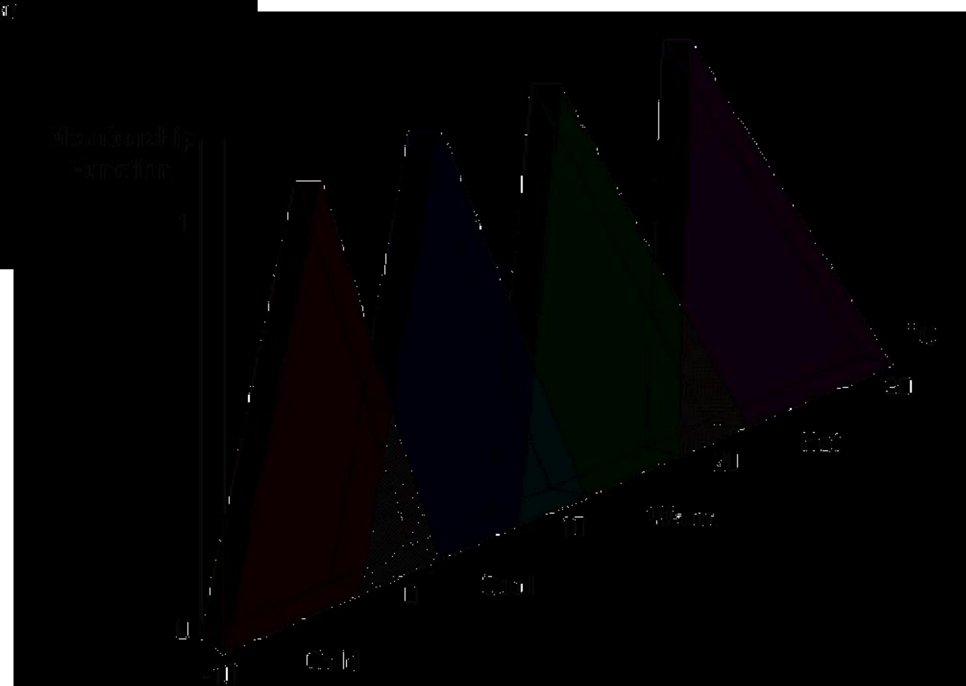

1.0
0.0
degree of
membership, μ
sharp-edged
membership
function for
TALL
height
tall (μ = 1.0)
not tall (μ = 0.0)


1.0
0.0
degree of
membership, μ
continuous
membership
function for
TALL
height
definitely a tall
person (μ = 0.95)
really not very
tall at all (μ = 0.30)

- Fuzziness is synonymous to inexactness.
- Difference between Probability and Fuzziness.
- When it is appropriate to use Fuzzy Logic?
- Fuzzy Logic and Fuzzy sets are based on the way the brain deals with inexact information.
- In precise, scientific tests and measurements also used “fuzzy” logic in intuitive manner to evaluate results, symptoms, relationships, causes, or remedies.
- Fuzzy systems afford a broader, richer field of data and manipulations than do more traditional methods.

Bivalence and Fuzzy





Basics of Crisp sets

- Generalization of Crisp Set.

Crisp set:

A *Crisp set* is a collection of distinct objects whose elements in a given universe of discourse are placed into two groups; *Members* and *nonmembers*.

Let U be a universe of discourse, the characteristics function $\mu_A(x)$ of a crisp set A in U takes its values in [0,1].

$$\mu_A(x) = 1 \text{ if and only if } x \in A$$
$$= 0 \text{ otherwise.}$$

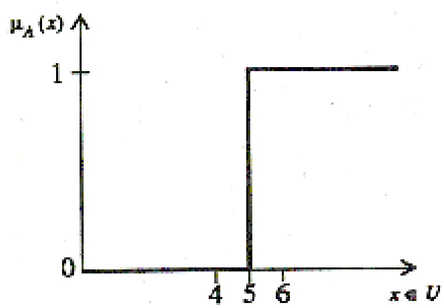
Example

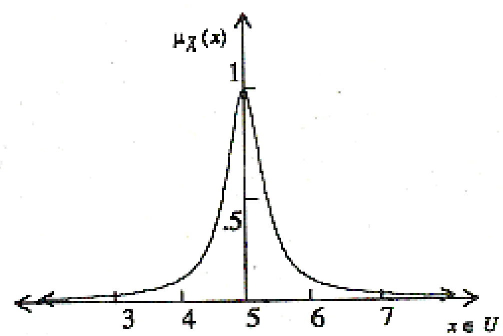
U be the real line $\Re$ and crisp set **A** represents “real numbers greater than and equal to 5”;

$$A=\{\ (x, \mu_A(x)) \mid x \in U\ \},\ \mu_A(x) \begin{aligned} &= 0 \quad x < 5 \\ &= 1, \quad x \geq 5 \end{aligned}$$

fuzzy set **A** represents “real numbers close to 5”;

$A = \{\ (x, \mu_A(x)) \mid x \in U\ \},\ \mu_A(x)=1/1+10(x-5)^2$



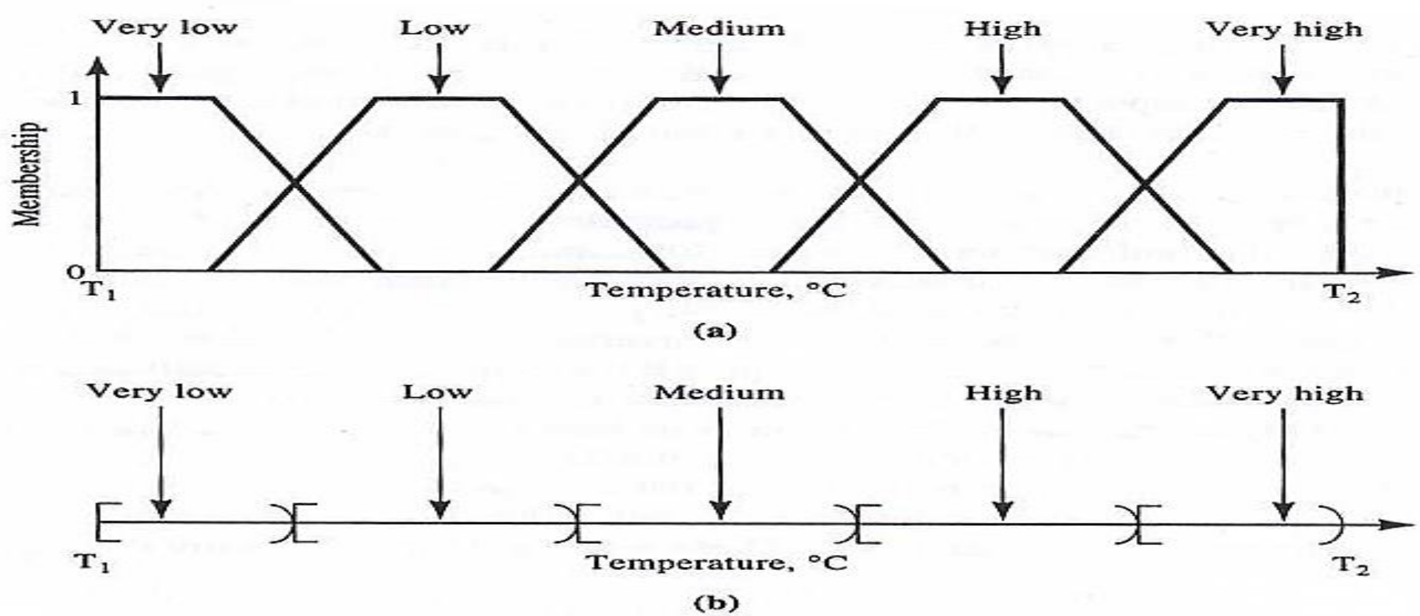


Membership function

A fuzzy set *A* in the universe of discourse *U* can be defined as a set of ordered pairs, $A = \{ (x, \mu_A(x)) \mid x \in U \}$

$\mu_A(x)$ is the grade of membership of x in A

- For fuzzy sets the range of membership function is [0,1].



Temperature in the range $[T_1, T_2]$ conceived as: (a) a fuzzy variable; (b) a traditional (crisp) variable.

Observations

- The boundary of ***crisp set*** is rigid and sharp and performs a classification into two classes (x € A or x$\notin$A)
- The universe of discourse **U** is a crisp set.
- A ***Fuzzy set*** introduces vagueness by eliminating the sharp boundary that divides members from nonmembers in a group.
- The transition between full membership and nonmembership is gradual rather than abrupt.
- *Fuzzy sets* may be viewed as an extension and generalization of *crisp sets*.

- The **crisp set** always has unique membership fn., whereas every **fuzzy set** has an infinite no. of membership fns. This enables *fuzzy systems* to be adjusted for maximum utility in a given situation.

- The membership fn. of a *fuzzy set* is subjective in nature; however it cannot be assigned arbitrarily.

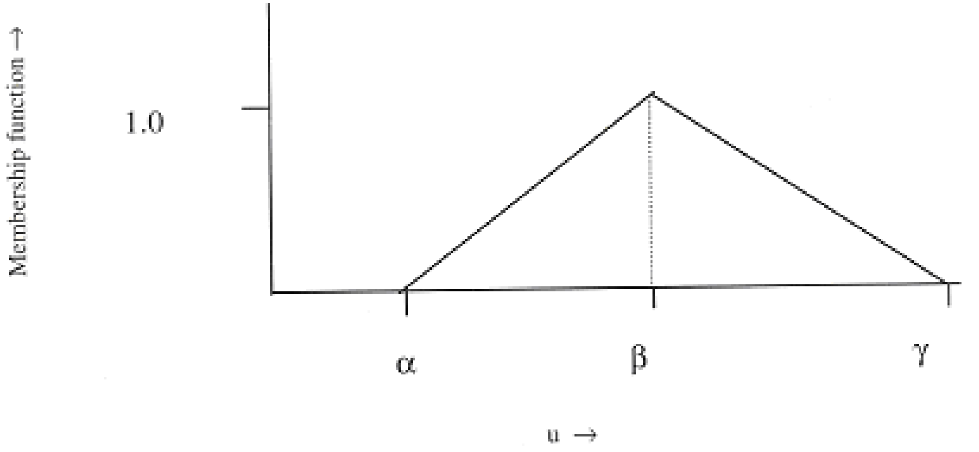
- Estimation of membership fn. is complicated and a better approach is to utilize the learning power of NNs to approximate them.

- In ***fuzzy*** $\sum\mu_A(x) \neq 1$ unlike probability.

Triangular Membership function

- The *triangular membership* fn. also called bell-shaped function with straight lines, can be defined as:

$$\Lambda(u;\alpha,\beta,\Upsilon) = \begin{cases} 0, & u \le \alpha \\ (u-\alpha)/(\beta-\alpha), & \alpha < u \le \beta \\ (\alpha-u)/(\beta-\alpha), & \beta < u \le \Upsilon \\ 0. & u > \Upsilon \end{cases}$$



Membership Functions

- A *Gaussian membership* function is defined as:

  G(u; m, σ) = exp [- {(u-m)/ 2σ}2],

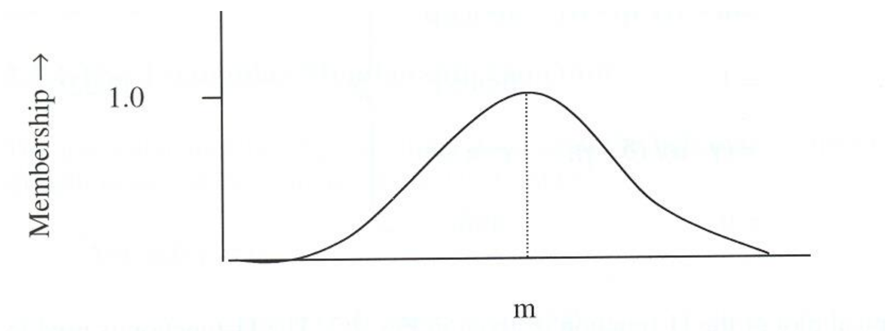
where *m* and *σ* control the mean and width of the membership function.

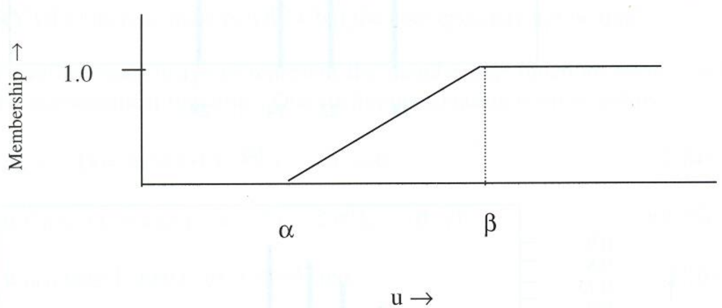
- ϒ membership function is defined as:

ϒ (u**;** α, β) = 0 ; u ≤ α

= (u-α)/(β-α) ; α < u ≤ β

= 1 **;** u > β





Basic Operations on Fuzzy Sets

- the *cardinality* (or *scalar cardinality)* of a fuzzy set A : $|\mathbf{A}| = \sum_{x \in U} \mu_A(x)$
- Equality: A and B are *equal* if and only if $\mu_A(x) = \mu_B(x)$, $\forall x \in U$
- To check the *degree of equality* of two fuzzy sets, we can use the *similarity measure:*
- $E(A, B) \equiv \text{degree}(A = B) = |A \cap B| \,/\, |A \cup B|$
- When $A = B$, $E(A, B) = 1$;
- when $|A \cap B| = 0$, $E(A, B) = 0$.
- In most general cases, $0 \leq E(A, B) \leq 1$.

Operations on Fuzzy sets

- *A* is a *subset* of *B;* if and only if $\mu_A(x) \leq \mu_B(x)$, $\forall x \in U$

- *subsethood measure* :S (*A*, *B*) = degree($A \subseteq B$) = $|A \cap B| \,/\, |A|$

- $\overline{A \cup B} = \overline{A} \cap \overline{B}$,

- $\overline{A} \cup A \neq U$ (*law of the excluded middle*) *and* $\overline{A} \cap A \neq \emptyset$ (*law of contradiction*)

Intersection: $\mu_{A \cap B}(x) = \min\,[\mu_A(x), \mu_B(x)] = \mu_A(x) \wedge \mu_B(x);$
$\forall x \in U$

Union: $\mu_{A \cup B}(x) = \max\,[\mu_A(x), \mu_B(x)] = \mu_A(x) \vee \mu_B(x);\ \forall x \in U$

- The Cartesian Product of $A_1, A_2, \ldots\ldots, A_n$ is a fuzzy set in the product space $U1 \times U2 \times \ldots. \times Un$ with membership function:

$\mu_{A1 \times A2 \times \ldots. \times An}(x_1, x_2, \ldots x_n) = \min\,[\mu_{A1}(x_1), \mu_{A2}(x_2), \ldots \mu_{An}(x_n)],$
$x_1 \in U1, x_2 \in U2, \ldots, x_n \in Un$

Algebraic sum: $\mu_{A+B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\ .\ \mu_B(x)$

Algebraic product: $\mu_{A.B}(x) = \mu_A(x)\ .\ \mu_B(x)$

Bounded sum: $\mu_{A \oplus B}(x) = \min\,\{1, \mu_A(x) + \mu_B(x)\}$

Bounded difference: $\mu_{A \Theta B}(x) = \max\,\{0, \mu_A(x) - \mu_B(x)\}$

Fuzzy T-Norm

the intersection of the fuzzy sets, characterized by a T-norm operator: $\mu_{A \cap B}(x) = T[\mu_A(x), \mu_B(x)]$

- for any membership values a, b, c and d, the T-norm operator :
- $T(0,0)=0$, $T(a,1)=T(1,a)=a$ (boundary)
- $T(a,b) \le T(c,d)$ if $a \le c$ and $b \le d$ (monotonicity)
- $T(a,b)=T(b,a)$ (commutativity)
- $T(a, T(b,c))=T(T(a,b),c)$ (associativity)
- **Def.:** A function $T:[0,1] \times [0,1] \rightarrow [0,1]$ that satisfies the above 4 characteristics is called a T-norm.

Fuzzy S-Norm

- The union of the fuzzy sets, characterized by a S-norm (T-co-norm) operator: $\mu_{A \cup B}(x) = S\,[\mu_A(x), \mu_B(x)]$
- $S(1,1) = 1$, $S(a, 0) = S(0, a) = a$ (boundary)
- $S(a, b) \leq S(c, d)$ if $a \leq c$ and $b \leq d$ (monotonicity)
- $S(a, b) = S(b, a)$ (commutativity)
- $S(a, S(b, c)) = S(S(a, b), c)$ (associativity)
- **Def.:** A function $S:[0,1] \times [0,1] \rightarrow [0,1]$ that satisfies the above 4 characteristics is called a S-norm.

FUZZY RELATION

Crisp Relation Set

- Relation is a new kind of crisp set expressed by rules:
- Rule 1: If the colour of the fruit is green then the fruit is verdant.
- Rule 2:If the colour of the fruit is yellow then the fruit is half-mature.
- Rule 3: If the colour of the fruit is red then the fruit is mature.

Crisp Relation

- Relationship between the colours of a fruit X and the grade of maturity Y

  X = {green, yellow, red}

  Y = {verdant, half-mature, mature}

- Crisp formulation of a relation X $\rightarrow$Y

	verdant	half-mature	mature
green	1	0	0
yellow	0	1	0
red	0	0	1

Crisp Relation

- Crisp relation is defined on the Cartesian product of two universal sets as: $X \times Y = \{(x, y) | x \in X, y \in \boldsymbol{Y}\}$
- The crisp relation R is defined by

$$\mu_R(x, y) = \{1, (x,y) \in R$$
$$= \{0, (x,y) \notin R$$

- "1" implies complete truth degree for the pair to be in relation and "0" implies no relation.
- When the sets are finite the relation R is called a relation matrix.
- Binary, Ternary, Quaternary, Quinary, n-ary relations

Crisp relation

- A *relation* among crisp sets $A_1,\ A_2,\ \ldots,\ A_n$ is a subset of the Cartesian product. It is denoted by R.

$$R \subseteq A_1 \times A_2 \times\ \ldots \times A_n$$

- Using the membership function defines the crisp relation R :

$$\mu_R(x_1,\ x_2,\ \ldots, x_n) = \begin{cases} 1 & \text{iff } (x_1,\ x_2,\ \ldots,\ x_n) \in R, \\ 0 & \text{otherwise} \end{cases}$$

where $x_1 \in A_1, x_2 \in A_2, \ldots, x_n \in A_n$

- Let $X = \{1,4,5\}$ and $Y = \{3,6,7\}$
- Classical matrix for the crisp relation $R:X<Y$

	3	6	7
1	1	1	1
4	0	1	1
5	0	1	1

Fuzzy Relation

- A *fuzzy relation* is a fuzzy set defined on the Cartesian product of crisp sets A_1, A_2, ..., A_n where tuples (x_1, x_2, ..., x_n) may have varying degrees of membership within the relation.

- The membership grade indicates the strength of the relation present between the elements of the tuple.

$$\mu_R : A_1 \times A_2 \times \ldots \times A_n \rightarrow [0,1]$$

$$R = \left\{((x_1, x_2, ..., x_n), \mu_R) \mid \mu_R(x_1, x_2, ..., x_n) \geq 0, x_1 \in A_1, x_2 \in A_2, ..., x_n \in A_n\right\}$$

Fuzzy Relation

- $\mathbf{X} = \{ x_1, x_2, \ldots..x_n \}$ and $\mathbf{Y}=\{y_1,y_2, \ldots.y_m \}$, R(**X**,**Y**) can be expressed by an $n \times m$ matrix as

$$R(X, Y) = \begin{bmatrix} \mu_R(x_1, y_1) & \mu_R(x_1, y_2) & \cdots & \mu_R(x_1, y_m) \\ \mu_R(x_2, y_1) & \mu_R(x_2, y_2) & \cdots & \mu_R(x_2, y_m) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_R(x_n, y_1) & \mu_R(x_n, y_2) & \cdots & \mu_R(x_n, y_m) \end{bmatrix}.$$

- **Ex.** $X=\{x_1, x_2, x_3\}$ and $Y=\{y_1, y_2\}$ and $R(x_i, y_j)$, represents the distance between x_i and y_j close to zero for all $x \in X$ and for all $y \in Y$.
- $R(x_i, y_j) = e^{[-(xi - yj)2]}$
- The expression describes "*X* close to Y".

Composition Operation

- Fuzzy relation R1 and R2 defined on X×Y, Y×Z, respectively.
- The **max-min composition** of R1 and R2 is a fuzzy set,

$$R3=R1\circ R2=\{\mu_{R3}(x,z)\}=\max_{y}\{\min(\mu_{R1}(x,y),\mu_{R2}(y,z))|\ x\in X,y\in Y,z\in Z$$

- The **min-max composition** of R1 and R2 is a fuzzy set,

$$R3=R1\circ R2=\{\mu_{R3}(x,z)\}=\min_{y}\{\max(\mu_{R1}(x,y),\mu_{R2}(y,z))|\ x\in X,y\in Y,z\in Z$$

- There is **max-min** and **min-max** composition for set-relation compositions.

- **A** be a fuzzy set on **X** and R(X,Y) is a fuzzy relation on X×Y,

$$\mu_{A\circ R}(y)=\max_{x\in X}\min[(\mu_A(x),\mu_R(x,y)];\quad \text{for all } y\in Y,$$

The max-min composition

- The max-min composition is commonly used when a system requires conservative solutions in the sense that the goodness of one value cannot compensate the badness of another value.
- Associative: $[P(X, Y) \circ Q(Y, Z)] \circ R(Z, W) =$
  $P(X, Y) \circ [Q(Y, Z) \circ R(Z, W)]$
- $[P(X, Y) \circ Q(Y, Z)]^{-1} = Q(Z, Y)^{-1} \circ P(Y, X)^{-1}$
- $P(X, Y) \circ Q(Y, Z) \neq Q(Y, Z) \circ P(X, Y)$

α-cut of fuzzy relation

0.9	0.4	0.0
0.2	1.0	0.4
0.0	0.7	1.0
0.4	0.2	0.0

$M_{R\,0.4} =$

1	1	0
0	1	1
0	1	1
1	0	0

$M_{R\,0.9} =$

1	0	0
0	1	0
0	0	1
0	0	0

$M_{R\,0.7} =$

1	0	0
0	1	0
0	1	1
0	0	0

$M_{R\,1.0} =$

0	0	0
0	1	0
0	0	1
0	0	0

α-cut of fuzzy relation

- Assume $R \subseteq A \times B$, and R_α is a α-cut relation. Then
  $$R_\alpha = \{(x, y) \mid \mu_R(x, y) \geq \alpha, \quad x \in A, y \in B\}$$
  Note that R_α is a crisp relation. □

- Example

$$M_R = \begin{array}{c|ccc} & & & \\ \hline & 0.9 & 0.4 & 0.0 \\ & 0.2 & 1.0 & 0.4 \\ & 0.0 & 0.7 & 1.0 \\ & 0.4 & 0.2 & 0.0 \end{array} \qquad \Lambda = \{0, 0.2, 0.4, 0.7, 0.9, 1.0\}$$

Decomposition of relation

Definition (Decomposition of relation) Fuzzy relation can be said to be composed of several R_α's as following.

$$R = \bigcup_{\alpha} \alpha R_\alpha$$

Here α is a value in the level set; R_α is a α-cut relation; αR_α is a fuzzy relation. The membership function of αR_α is defined as,

$$\mu_{\alpha R_\alpha}(x, y) = \alpha \bullet \mu_{R\alpha}(x, y) \quad , \quad \text{for } (x, y) \in A \times B$$

Thus we can decompose a fuzzy relation R into several αR_α. □

Decomposition of relation

Example 3.8 The relation R in the previous example can be decomposed as following.

$$M_R = 0.4 \times \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{vmatrix} \cup 0.7 \times \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{vmatrix} \cup 0.9 \times \begin{vmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} \cup 1.0 \times \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix}$$

$$M_R = \begin{vmatrix} 0.9 & 0.4 & 0.0 \\ 0.0 & 1.0 & 0.4 \\ 0.0 & 0.7 & 1.0 \\ 0.4 & 0.0 & 0.0 \end{vmatrix}$$

Operations on fuzzy relations

- Complement relation:

$$\forall (x, y) \in A \times B$$
$$\mu_{\overline{R}}(x, y) = 1 - \mu_R(x, y)$$

- Example

M_R	a	b	c
1	0.3	0.2	1.0
2	0.8	1.0	1.0
3	0.0	1.0	0.0

$M_{\overline{R}}$	a	b	c
1	0.7	0.8	0.0
2	0.2	0.0	0.0
3	1.0	0.0	1.0

Operations on fuzzy matrices

- Sum: $A + B = \max[a_{ij}, b_{ij}]$

- Example

$A =$

	a	b	c
a	0.2	0.5	0.0
b	0.4	1.0	0.1
c	0.0	1.0	0.0

$B =$

	a	b	c
a	1.0	0.1	0.0
b	0.0	0.0	0.5
c	0.0	1.0	0.1

$A + B =$

	a	b	c
a	1.0	0.5	0.0
b	0.4	1.0	0.5
c	0.0	1.0	0.1

Fuzzy Implication Relations

If x is A1(p) Then y is B1(q) is denoted by: $R1(x,y)=\{\mu_{R1}(x,y)/(x,y)\}$

- $\mu_{R1}(x,y)$ is constructed by

(i) **Dienes-Rescher Implication:** $Max[1- \mu_{A1}(x),\mu_{B1}(y)]$; $p\rightarrow q:\sim p \vee q$

(ii) **Lukasiewicz Implication**: $Min[1,1- \mu_{A1}(x)+ \mu_{B1}(y)]$; Replace negation by 1's complement and logical OR by '+'.

(iii) **Mamdani Implication:** $Min[\mu_{A1}(x), \mu_{B1}(y)]$ or $\mu_{A1}(x).\mu_{B1}(y)$

(iv) **Zadeh Implication:** $Max[Min[\mu_{A1}(x), \mu_{B1}(y), 1- \mu_{A1}(x)]$; $p\rightarrow q$ can be stated as "either p and q are true or p is false" $(p\wedge q)\vee(\sim p)$.

(v) **Godel Implication:** $\mu_{R1}(x,y)=1$ if $\mu_{A1}(x)\leq \mu_{B1}(y)$
$= \mu_{B1}(y)$ otherwise

Binary Fuzzy Relations

- Binary Fuzzy relations on a single set X.
- Different types of binary fuzzy relations are distinguished by three basic properties:

reflexivity, Symmetry and Transitivity.

- A fuzzy relation R(X,X) is **reflexive** ; iff $\mu_R(x, x)=1; x \in X$.
- R(X,X) is **symmetric**; iff $\mu_R(x,y)= \mu_R(y,x)$; for all $x,y \in X$.
- R(X,X) is **Transitive;** iff $\mu_R(x,y) \geq \max\limits_{y \in Y} \min[\mu_R(x,y), \mu_R(y,z)]$;

$$\forall (x,z) \in X^2$$

Similarity Relation

- A binary fuzzy relation which is reflexive, symmetric and transitive binary fuzzy relation is called **Similarity Relation**.
- Similarity relations are the generalization of equivalence relations in binary crisp relations to binary fuzzy relations.
- A binary crisp relation C(X,X) is called an equivalence relation if it satisfies Reflexivity: $(x, x) \in C(X, X)$ for every $x \in X$
- Symmetry: If $(x,y) \in C(X, X)$ then $(y, x) \in C(X, X)$, $x,y \in X$
- Transitivity: : If $(x,y) \in C(X, X)$ and $(y, z) \in C(X, X)$, then $(x,z) \in C(X,X)$, $x,y,z \in X$

Fuzzy Measures

- In Fuzzy sets, a value is assigned to each element x of the universal set X signifying its degree of membership in a particular set with unsharp boundaries (fuzzy sets) since it is not possible to draw crisp boundary in deciding whether a person is *tall*.
- In some situation, like judging a person is guilty or not we need sharp boundary.
- In trail case, group of people as guilty and innocent are crisp sets.
- Evidence for trial judgment is rarely perfect and uncertainty prevails.

Fuzzy Measures

- This kind of uncertainty, where classes are defined with sharp boundaries (crisp set), but measurement is rarely perfect, creates problem in decision making.

- To represent this kind of uncertainty, known as ***ambiguity*****,** a value is assigned in unit interval [0,1] to each possible crisp set to which the element in question might belong.

- This value represents the degree of evidence or belief or certainty of the element's membership in the set, known as *fuzzy measures*.

Fuzzy Measures

- A fuzzy measure assigns a value in the unit interval [0,1] to each crisp set of the universal set signifying the degree of evidence or belief that a particular element x belongs in the crisp set.
- Several different measures, like belief measures, possibility measures, necessity measures and plausibility measures are applied to crisp subsets instead of elements, of a universal set.
- FS are used to solve vagueness associated with the difficulty of making sharp distinctions of objects in the universe.
- Fuzzy measures are used to solve ambiguity associated with making a choice between to or more alternatives.

Example: Consider a group of people X. For FSs, the age of a person $x \in X$ is known and consider x to be **old** with a membership grade $\mu_A(x)$, where A is a FS of **old** people.

- For FMs, the age of the person $x \in X$ is unknown and by looking at that person, we can consider the person to be in the crisp set A consisting of people who are more than 50 years old with a measure of $g_x(A)$.

- $g_x(A)$ is a FM that assigns a value to each possible crisp set (each age group of people) to which the person in question might belong.

Fuzzy Measures

- So, in FSs, a value is assigned to each element of the universal set signifying its degree of membership in a particular set with an unsharp boundary.

- While in FMs, a value is assigned to each crisp set of the universal set signifying the degree of evidence or belief that a particular element belongs in the set.

- It is to be noted that the element x in FS A has been located to A, while the element x in a FM has not been previously located to the crisp set A.

- A fuzzy measure is defined by a set function

$$g: \wp(X) \rightarrow [0,1]$$

which assigns to each crisp subset of a universe of discourse X a number in the unit interval [0,1], where $\wp(X)$ is the power set of X.

- When this number is assigned to a crisp subset $A \in \wp(X)$, g(A) represents the degree of evidence or belief that a given element $x \in X$ (which has not been previously located in any crisp subset of X) belongs to the crisp subset A.

- The domain of the function *g* is the power set $\wp(X)$ of crisp subsets of X and not the power set $\wp(X)$ of fuzzy subsets of X.

- A fuzzy measure is a set function must have certain properties.

- A fuzzy measure is a set function g: $\Im \rightarrow [0,1]$ where $\Im \subset \wp(X)$ is a family of crisp subsets of X and is a Borel field and *g* satisfies the axioms.

- Axiom *g*1 (Boundary conditions): $g(\phi) = 0$ and $g(X) = 1$;

  Element in question definitely does not belong to the empty set and definitely belongs to the universal set.

- Axiom *g*2 (Monotonicity): For every crisp set A, B $\in \wp(X)$, if A $\subseteq$ B, then $g(A) \leq g(B)$.

- Axiom g3 (continuity): For every sequence $(A_i \in \wp(X) \mid i \in \aleph)$ of subsets of X, if either $A_1 \subseteq A_2 \subseteq \ldots\ldots$ Or $A_1 \supseteq A_2 \supseteq \ldots$(the sequence is monotonic) then $\lim g(A_i) = g(\lim A_i)$; $i \rightarrow \infty$

where $\aleph$ is the set of all positive integers.

- Axiom g1 states that we always know that the element in question definitely does not belong to the empty set and definitely does belong to the universal set.
- Axiom g2 requires that the degree of evidence of an element in a set (i.e. B) must be at least as great as the evidence that the element belongs to any subset of that set.
- Axiom g3 is applicable only to an infinite universal set and g is a continuous function.
- The concept of fuzzy measure was introduced by Sugeno to exclude the additivity requirement of standard measures, h.

- The additivity requirement of standard measure h indicates that when two sets A and B are disjoint ($A \cap B = \varnothing$) then $h(A \cup B) = h(A) + h(B)$

- Probability measure follows additivity property but fuzzy measures are defined by subadditivity, thus probability measure a special type of fuzzy measure.

- $\max[g(A), g(B)] \leq g(A \cup B)$; as $A \subseteq A \cup B$, $B \subseteq A \cup B$ and monotonic property of the FM g.

$g(A \cap B) \leq \min[g(A), g(B)]$; as $A \cap B \subseteq A$ and $A \cap B \subseteq B$

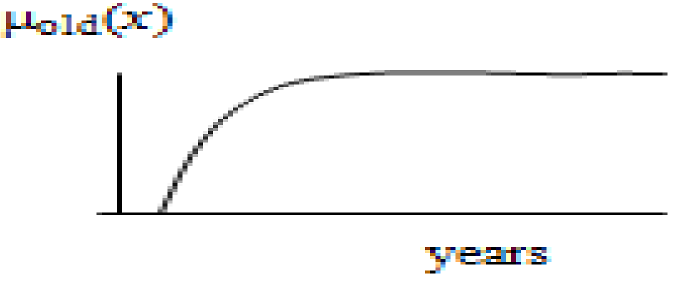
Fuzzy Logic

- Fuzzy Logic is an extension of Classical two-valued logic.
- The degree of an element in a fuzzy set may corresponds to the truth value of a proposition in Fuzzy Logic.
- Linguistic variables and Approximate Reasoning are two important concepts in fuzzy logic.
- A Linguistic variable is a variable whose values are words or sentences in a natural language.
- *Temperature* is a linguistic variable if it takes the values such *low*, *high*, *very high*
- The concept of linguistic variables provide a means of approximate characterization of phenomena that are too complex or too ill defined to be amanable to description in conventional quantitative terms.

- Linguistic variables and linguistic truth value is associated with fuzzy proposition.

- Fuzzy variable: (X,U,R(X)); X: name of the variable, R(X): A fuzzy subset of U, represents a fuzzy restriction imposed by X.

- Ex. X= ‘old’ ; U={10,20, …., 80};
  R(X)=0.1/20+0.2/30+0.4/40+0.5/50+0.8/60+1/70+1/80



A linguistic variable is a variable of higher order than a fuzzy variable and takes fuzzy variables as its values.

Linguistic Variables

linguistic variable: ”speed”; ***value***: ‘slow’, ‘fast’, ‘very fast’.

- Linguistic variable : (x, T(x), U, G, M)
- x: name of the variable; “speed”; T(x): term set of x;
- T(speed)={very slow, slow, moderate, fast …..}
- U=[0,100];
- G: syntactic rule for generating the names of the elements in T(speed), is quite intuitive.
- M=semantic rule; M(slow): the fuzzy set for “a speed below about 40 miles per hour” with membership function μ_{slow}

- Fuzzy Proposition: Fuzzy Predicates, fuzzy predicate modifiers, fuzzy quantifiers and fuzzy qualifiers.

- “Mary is Young”; “young” is fuzzy predicate.

- “The house is extremely expensive”; “extremely” is fuzzy predicate modifier.

- ”Many students are happy”; “Many” is fuzzy quantifier.

- (John is old) is Not Very True; “Not Very True” Fuzzy qualifier.

Drawback of Production System

- IF (Height is Tall) THEN (Speed is High) , KB: ”Height is VERY TALL”,
- No exact match occurs with the antecedent.
- In FL rule fires, ***inference***: “Speed is Very High”.
- **Extension in Fuzzy Domain:** IF (a banana is YELLOW) THEN (it is RIPE).

  VERY YELLOW, MORE-OR-LESS YELLOW, RIPE, VERY RIPE and MORE-OR-LESS-RIPE.

- Firing of these rules happen concurrently: “color” of the banana.
- “The banana is Yellowish” semantically close to “YELLOW”, “VERY YELLOW”…..
- “MORE-OR-LESS TRUE” or “ALMOST TRUE” is decided as “TRUE” in FL.

Linguistic Approximation

- Linguistic approximation is a procedure for determining a term from the term set of a linguistic variable which is semantically closest to a given fuzzy set.

- Linguistic approximation based on Similarity Measures.

• Linguistic approximation of a given fuzzy set A is the term $T_A \in T(x)$, which is the most similar to A compared to other terms in the term set $T(x)$ of a given linguistic variable.

$$E(A, T_A) = \max_{T_i \in T(x)} E(A, T_i)$$

- Given U = {0, 0.1, 0.2, …, 1} and the term set {A, B, C} of the linguistic variable *truth* where
- A = “True” = [{0.7,0.8}, {1,0.9}, {1,1}]
- B =“More or Less True”=[{0.5,0.6},{0.7,0.7},{1,0.8},{1,0.9}, {1,1}]
- C = “Almost True” = [{0.6,0.8}, {1,0.9},{0.6,1}]

- Find the linguistic approximation of the fuzzy set D defined by- D = [{0.6,0.8}, {1,0.9}, {1,1}]
- E(D,A) = 2.6/2.7 = 0.96, E(D,B) = 2.6/4.2 = 0.62,

  E(D,C) = 2.2/2.6 = 0.85

- The linguistic approximation of the fuzzy set D is the term A, “True”

Propositions in Logic

- Classical bivalence logic deals with *propositions* having value either *true* (logic value 1) or *false* (logic value 0).
- Propositions are sentences expressed in some language and expressed in a *canonical form:*- *x* is P, where *x* is a symbol of the *subject* and P designates the *predicate* which characterizes a property of the subject.
- A proposition and its negation are required to assume opposite truth values.
- Two concerns of logic: *logic operations* and *logic reasoning*

Logic Operations

- Logic operations are logic functions of two propositions and defined by truth values.

- Consider two propositions A and B, either of which can be *true* or *false*.

- Basic logic operations are conjunction (A *and* B), disjunction (A *or* B), implication (*if* A *then* B) and equivalence (*A if and only if* B)

Logic is a basis for Reasoning

Inference Rules:

- Modus Ponens: $(A \wedge (A \to B) \Rightarrow B$, i.e. *inference* B.
- Modus Tollens: $\sim B \wedge (A \to B) \Rightarrow \sim A$, i.e. *inference* $\sim A$.
- Hypothetical Syllogism: $(A \to B) \wedge (B \to C) \Rightarrow A \to C$
- The inferred propositions are always true no matter what the truth values of propositions A and B, called *tautologies*.
- Modus ponens are related to *Forward data driven inference* and useful for fuzzy logic control.
- *Modus tollens is* closely related to the *backward goal driven inference* usually used in Expert Systems

- **GENERALIZED MODUS PONEN**:

IF *x* is A THEN *y* is B; Given: *x* is A′; Inferred: *y* is B′

$$\mu_{B'}(y) = \mu_{A'}(x) \circ \mu_R(x, y)$$

- **GENERALIZED MODUS TOLLEN:**

IF *x* is A THEN *y* is B; Given: *y* is B′; Inferred: *x* is A′

$$\mu_{A'}(x) = \mu_{B'}(y) \circ [\mu_R(x, y)]^T$$

- Fuzzy Logic is an extension of set -theoretic Bivalence Logic in which the truth values are terms of the linguistic variable *truth*

Definition of fuzzy logic:

A form of knowledge representation suitable for notions that cannot be defined precisely, but which depend upon their contexts.

Approximate Reasoning

- The final conclusion is *approximate rather than exact*.

- Multiple Antecedent clauses:

If (x is A) AND (y is B) Then (z is C); Where $x \in X$, $y \in Y$ and $z \in Z$

$T[\mu_A(x), \mu_B(y)]$ = antecedent membership distribution.

$\mu_R(x, y; z)$:implication rule; $\mu_{A'}(x)$ and $\mu_{B'}(y)$ are supplied. Find out $\mu_{C'}(z)$.

$\mu_{C'}(z) = t(\mu_{A'}(x), \mu_{B'}(y)) \circ \mu_R(x, y; z)$

IF (height is TALL) and (weight is MODERATE) THEN (Speed is HIGH).

Approximate Reasoning

$\mu_{TALL}(height)=\{0.5/5',0.8/6',1.0/7'\};$

$\mu_{MODERATE}(weight)=\{0.7/45kg,0.9/50kg\}$

$\mu_{HIGH}(speed)=\{0.6/6m/s, 0.8/8m/s, 0.5/9m/s\}$

Given:$\mu_{TALL'}(height)=\{0.6/5',0.7/6',0.9/7'\};$

$\mu_{MODERATE'}(weight)=\{0.8/45kg,0.7/50kg\}$

Find out: $\mu_{HIGH'}(speed)$

Assume TALL' = TALL, MODERATE' = MODERATE and HIGH' = HIGH

Antecedent membership distribution = $T(\mu_{TALL}(\text{height}), \mu_{MODERATE}(\text{weight}))$
$=\{(0.5 \wedge 0.7/5', 45\text{kg}), (0.5 \wedge 0.9/5', 50\text{kg}), \ldots\ldots\ldots\ldots\}$

μ_R (height, weight; speed): implication rule.

Lukasiewicz function: $\mu_R(5', 45\text{kg}; 6\text{m/s}) = \text{Min}[1, 1-0.5+0.6] = 1.0$

$$
\begin{array}{c c}
 & \begin{array}{ccc} 6\text{m/s} & 8\text{m/s} & 9\text{m/s} \end{array} \\
\begin{array}{c} 5',45\text{kg} \\ 5',50\text{kg} \\ 6',45\text{kg} \\ 6',50\text{kg} \\ 7',45\text{kg} \\ 7',50\text{kg} \end{array} &
\begin{pmatrix} 1 & . & . \\ & & \\ & & \\ & & \\ & & \\ & & \end{pmatrix}
\end{array}
$$

$\mu_{HIGH'}(\text{speed}) =$

$T(\mu_{TALL'}(\text{height}), \mu_{MODERATE'}(\text{weight})) \circ \mu_R(\text{height}, \text{weight}; \text{speed})$

Fuzzy Inference Systems

- A Fuzzy Inference System (FIS) is a way of mapping an input space to an output space using fuzzy logic.

- FIS uses fuzzy membership functions and rules.

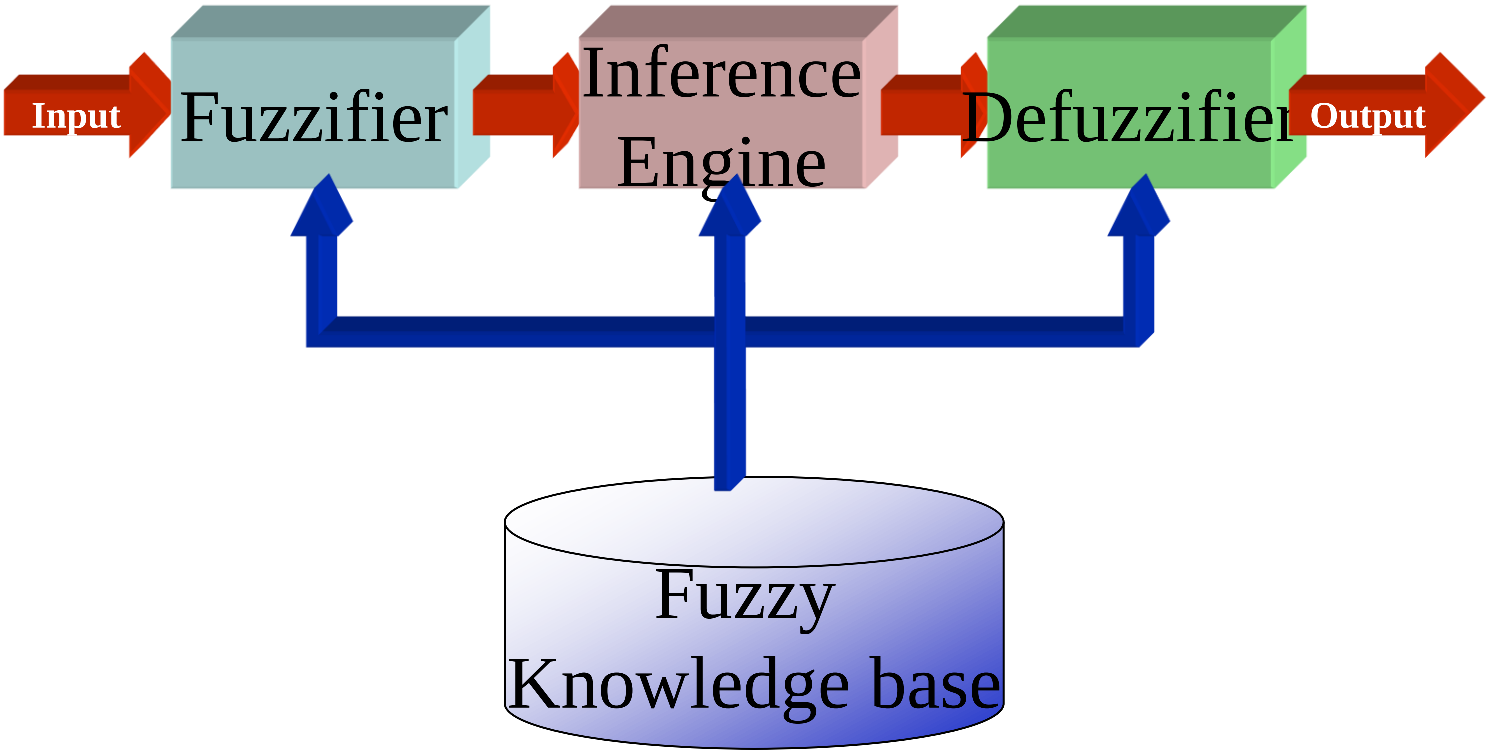
- The rules in FIS are fuzzy rules of the form:

  - *if p then q*, where *p* and *q* are fuzzy propositions.

- For example, in a fuzzy rule
  - if *x* is low and *y* is high then *z* is medium.
  - *x* is low; *y* is high; *z* is medium are fuzzy propositions
  - *x* and *y* are input linguistic variables; *z* is an output linguistic variable, which are crisps
  - low, high, and medium are fuzzy sets.

- The antecedent describes to what degree the rule applies, while the conclusion assigns a fuzzy function to each of one or more output variables.
- The set of rules in a fuzzy expert system is known as *knowledge base*.
- The functional operations in fuzzy expert system proceed in the following steps.
  - Fuzzification
  - Fuzzy Inferencing (apply implication method)
  - Aggregation of all outputs
  - Defuzzification

Fuzzy Inference Systems



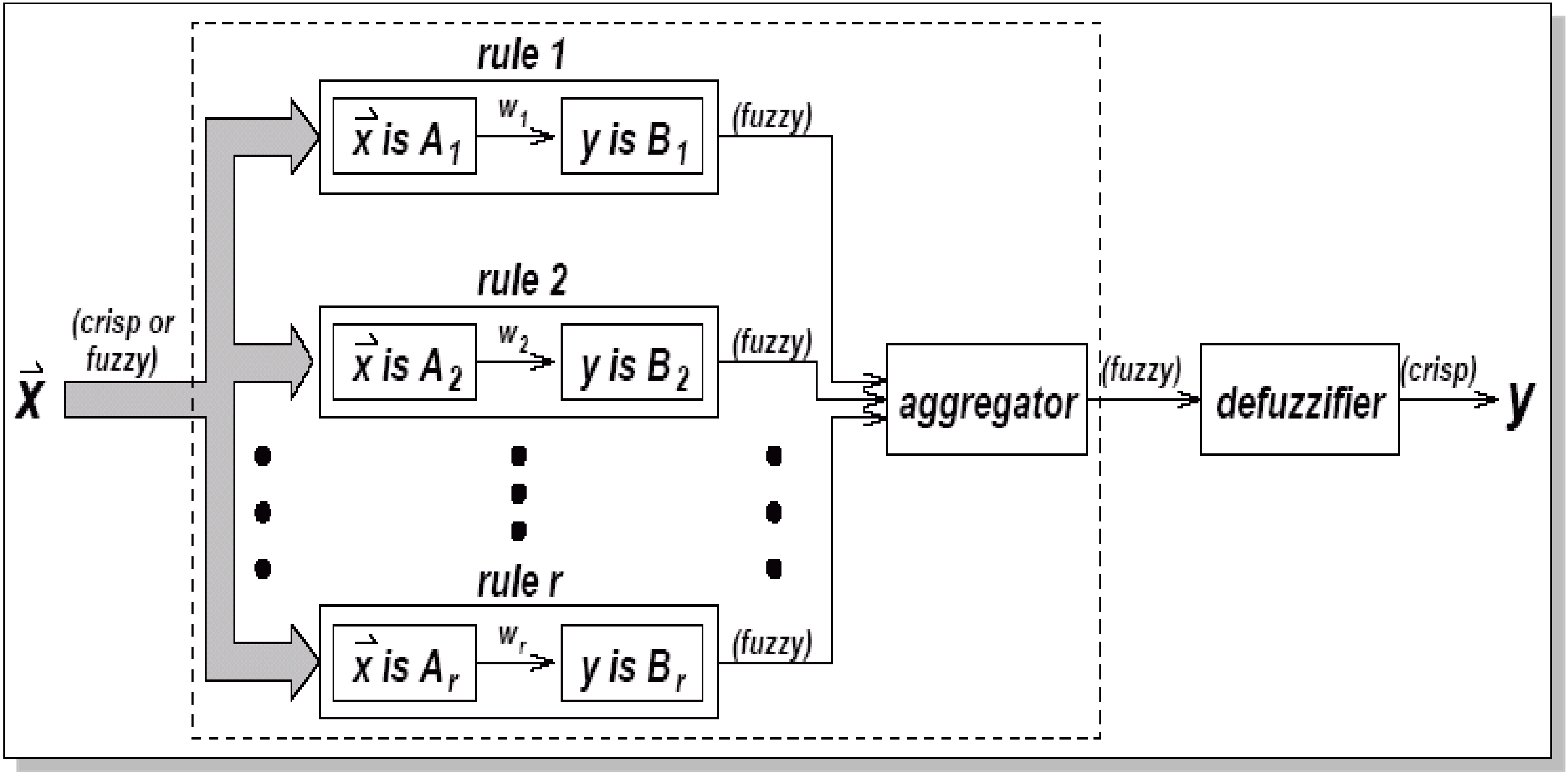
Fuzzification

- In fuzzification process, membership functions defined on input variables are applied to their actual values so that the degree of truth for each rule premise can be determined.

- Fuzzy statements in the antecedent are resolved to a degree of membership.
  - If there is only one part to the antecedent, then this is the degree of support for the rule.
  - If there are multiple parts to the antecedent, apply fuzzy logic operators and resolve the antecedent to a single number between 0 and 1.

- Antecedent may be joined by OR; AND operators.
  - For OR -- max
  - For AND -- min

Fuzzy Expert System



Fuzzy Inferencing

- The process of inference for each rule:
  - Truth value for the premise of each rule is computed and applied to the conclusion part of each rule.
  - A fuzzy set is assigned to each output variable in each rule.
- The use of degree of support for the entire rule shape the output fuzzy set.
- Combining the fuzzified inputs according to the fuzzy rules to establish a rule strength
- Finding the consequence of the rule by combining the rule strength and the output membership function.

- If the consequent of a rule has multiple parts, then all consequents are affected equally by the result of the antecedent.

- The consequent specifies a fuzzy set to be assigned to the output.

- The implication function then modifies that fuzzy set to the degree specified by the antecedent.

- Functions are used in inference rules.
- *min* or *prod* are commonly used as inference rules.
  - *min*: truncates the consequent's membership function
  - *prod*: scales it.

Aggregation of outputs

- Outputs of each rule are combined into a single fuzzy set.
- The output of the aggregation process is one fuzzy set for each output variable using a fuzzy aggregation operator.
- Some of the most commonly used aggregation operators are
  - the maximum : point-wise maximum over all of the fuzzy sets
  - the sum : (point-wise sum over all of the fuzzy
  - the probabilistic sum.

Fuzzy Inference Methods

Two types of fuzzy inference methods are **Mamdani** and **Sugeno-**

Mamdani fuzzy inference, introduced by Mamdani and Assilian (1975).

Sugeno or Takagi–Sugeno–Kang method introduced by Sugeno (1985).

The main difference between the two methods lies in the **consequent of fuzzy rules**.

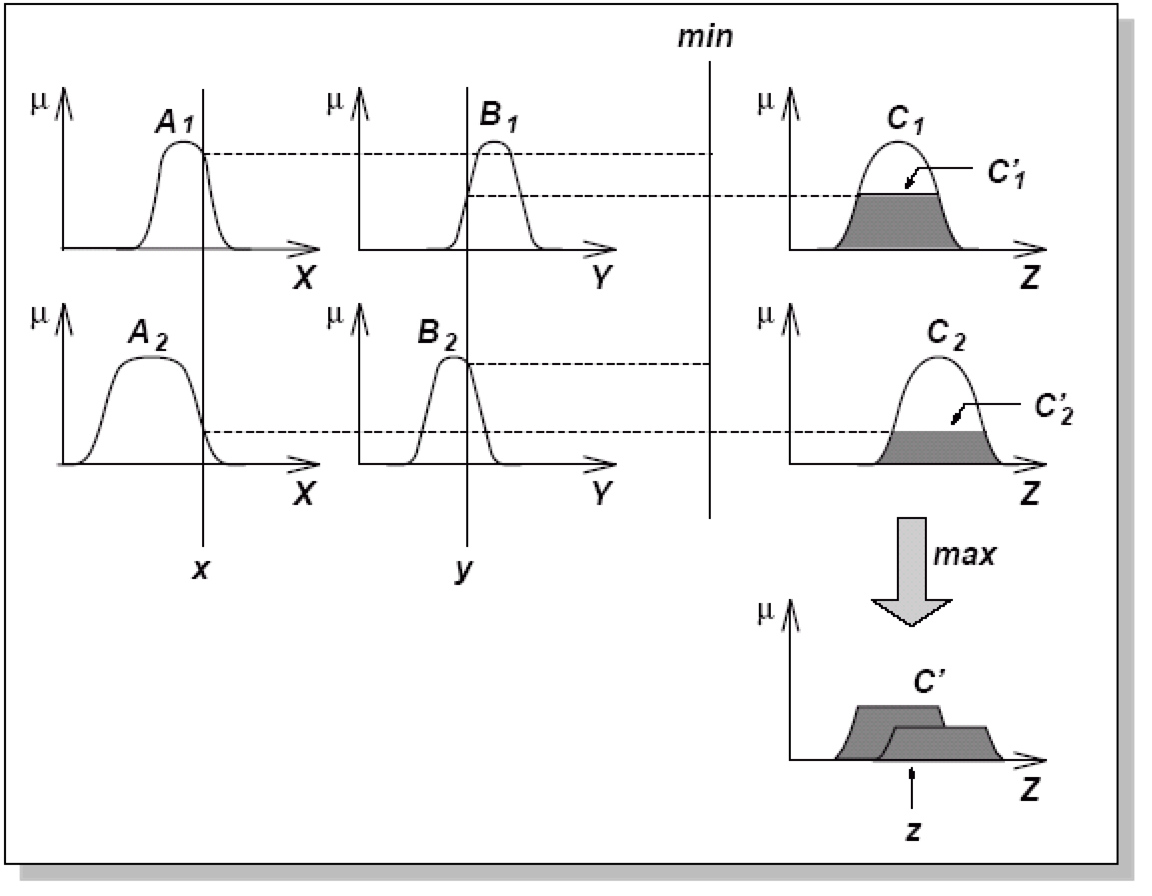
Mamdani Fuzzy models

compute the output of FIS given the inputs using steps below-

1. Determining **a set of fuzzy rules**
2. **Fuzzifying** the inputs using the input membership functions
3. **Combining the fuzzified inputs** according to the fuzzy rules to establish a **rule strength** (Fuzzy Operations)
4. Finding the **consequence of the rule** by combining the rule strength and the output membership function (implication)
5. **Combining the consequences** to get an output distribution (aggregation)
6. **Defuzzifying** the output distribution (this step is only if a crisp output (class) is needed).

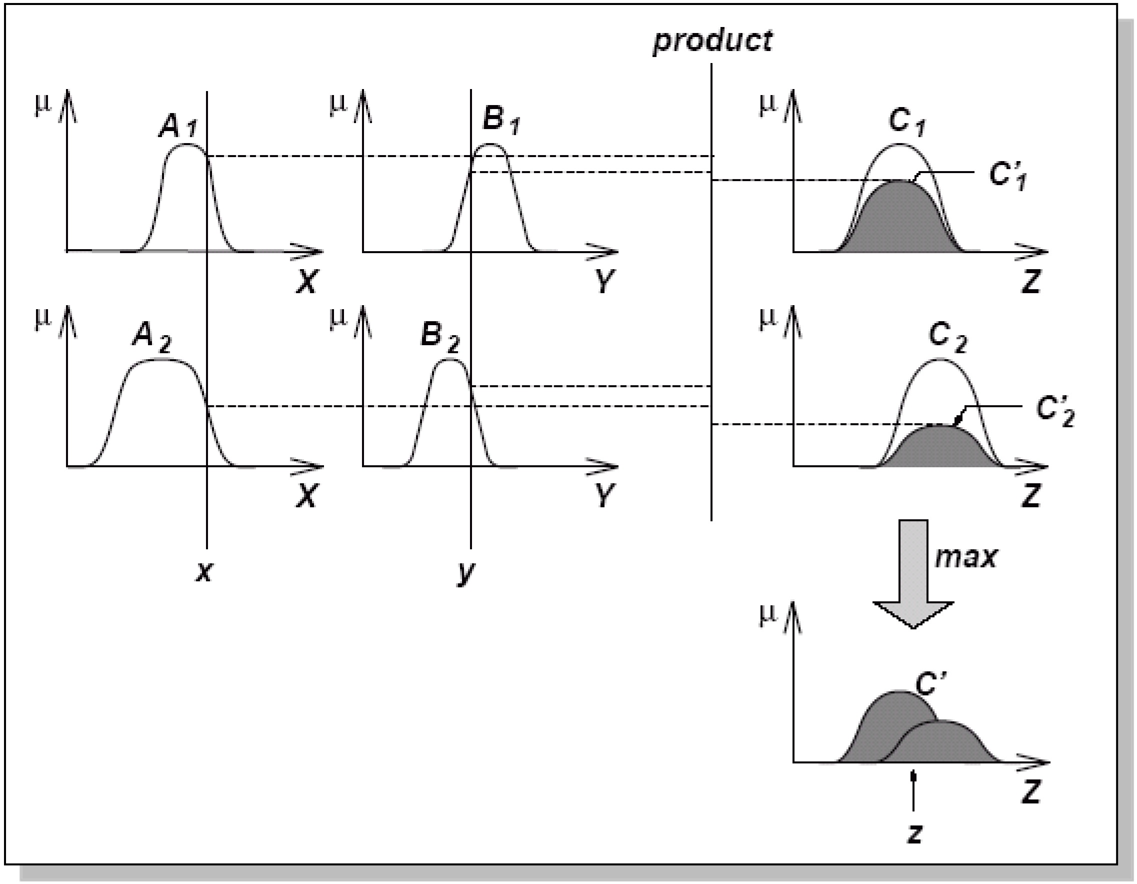
The Reasoning Scheme

Max-Min Composition is used



The Reasoning Scheme

Max-Product Composition is used.



Sugeno Fuzzy Models

- Also known as TSK fuzzy model
  - Takagi, Sugeno & Kang

- Goal: Generation of fuzzy rules from a given input-output data set.

Generic Method

- Steps :
  - Evaluate the antecedent for each rule
  - Obtain each rule's conclusion
  - Aggregate conclusions
  - Defuzzification

Examples

R1: if X is small and Y is small then $z = -x + y + 1$

R2: if X is small and Y is large then $z = -y + 3$

R3: if X is large and Y is small then $z = -x + 3$

R4: if X is large and Y is large then $z = x + y + 2$

Fuzzy Rules of TSK Model

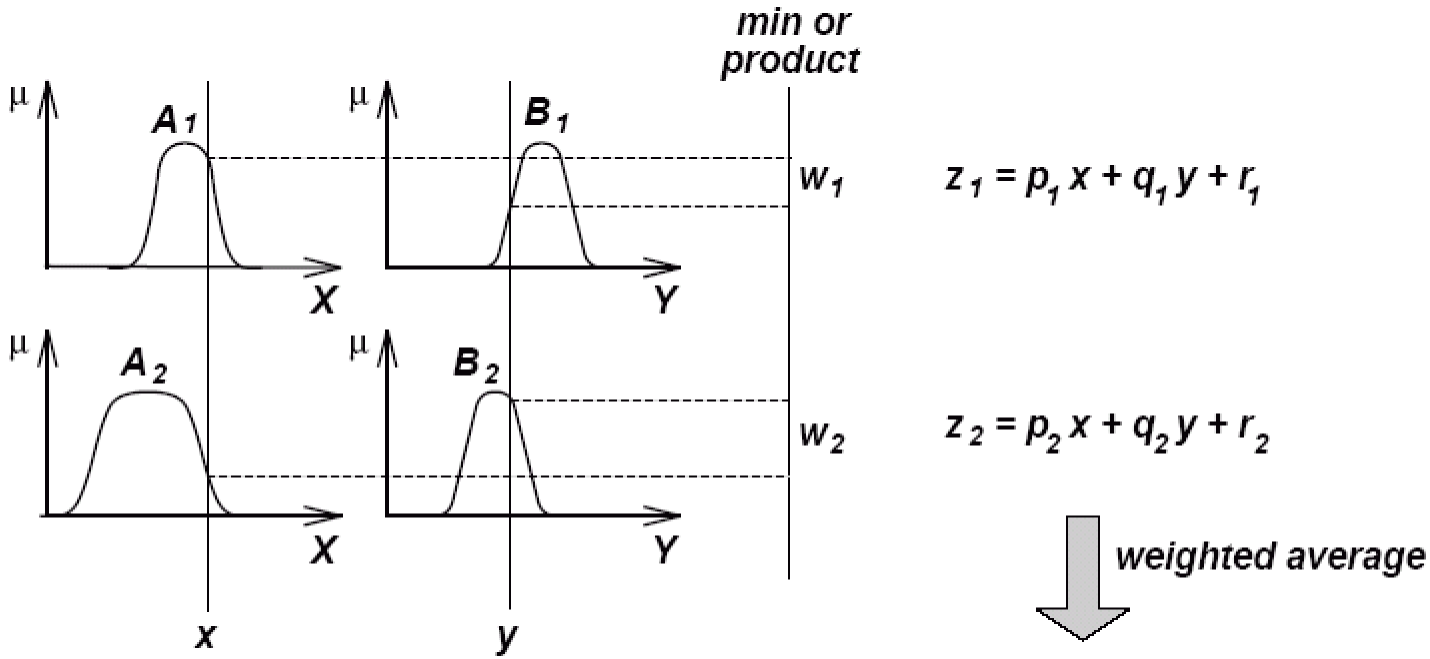
If x is A and y is B then $z = f(x, y)$

Fuzzy Sets

Crisp Function

$f(x, y)$ is very often a polynomial function

The Reasoning Scheme



$$z_1 = p_1 x + q_1 y + r_1$$

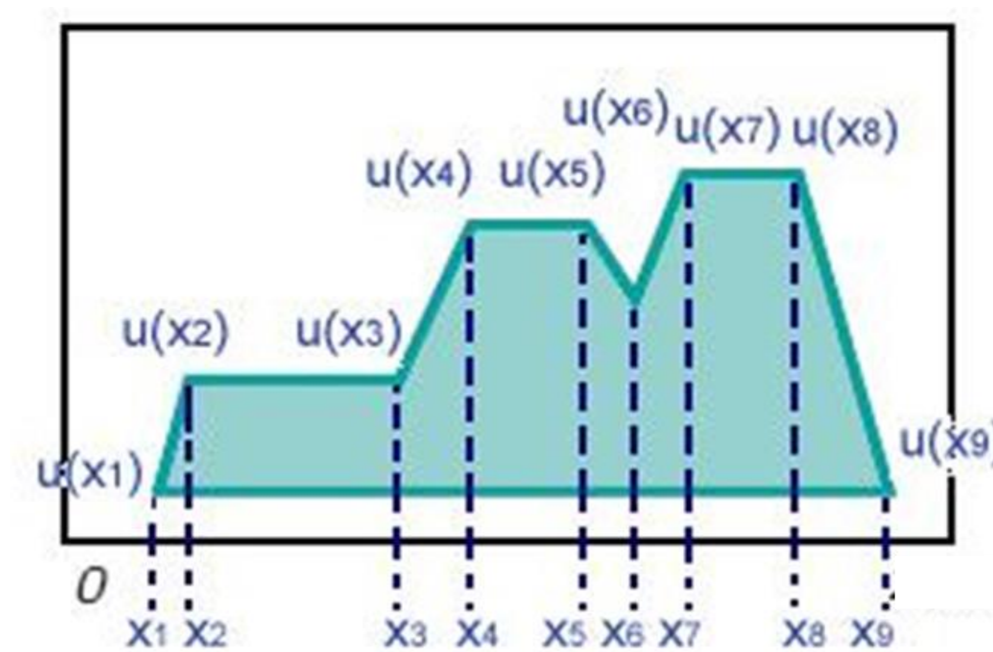
$$z_2 = p_2 x + q_2 y + r_2$$

$$z = \frac{w_1 z_1 + w_2 z_2}{w_1 + w_2}$$

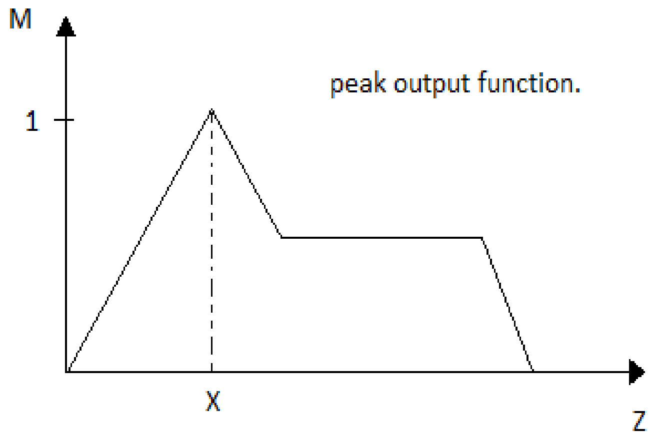
Defuzzification

- In Defuzzificztion, the fuzzy output set is converted to a crisp number.
- Techniques: *centroid*, *weighted average* and *maximum* methods.

$$g = \frac{\sum_{i=1}^{9} x_i \cdot u(x_i)}{\sum_{i=1}^{9} u(x_i)}$$



Centroid Method



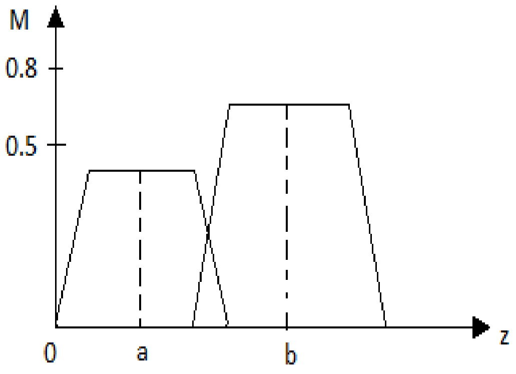
Maximum Method

Each membership function is weighted by its max membership value.

$$X^* = \frac{\sum Mc(\bar{xi}).\bar{xp}}{\sum MC(\bar{xi})}$$

$\bar{X}i$ = maximum of with member function.

$\sum$ = algebraic sum.



$$x^* = \frac{0.5a+0.8b}{0.5+0.8}$$

Weighted Average method

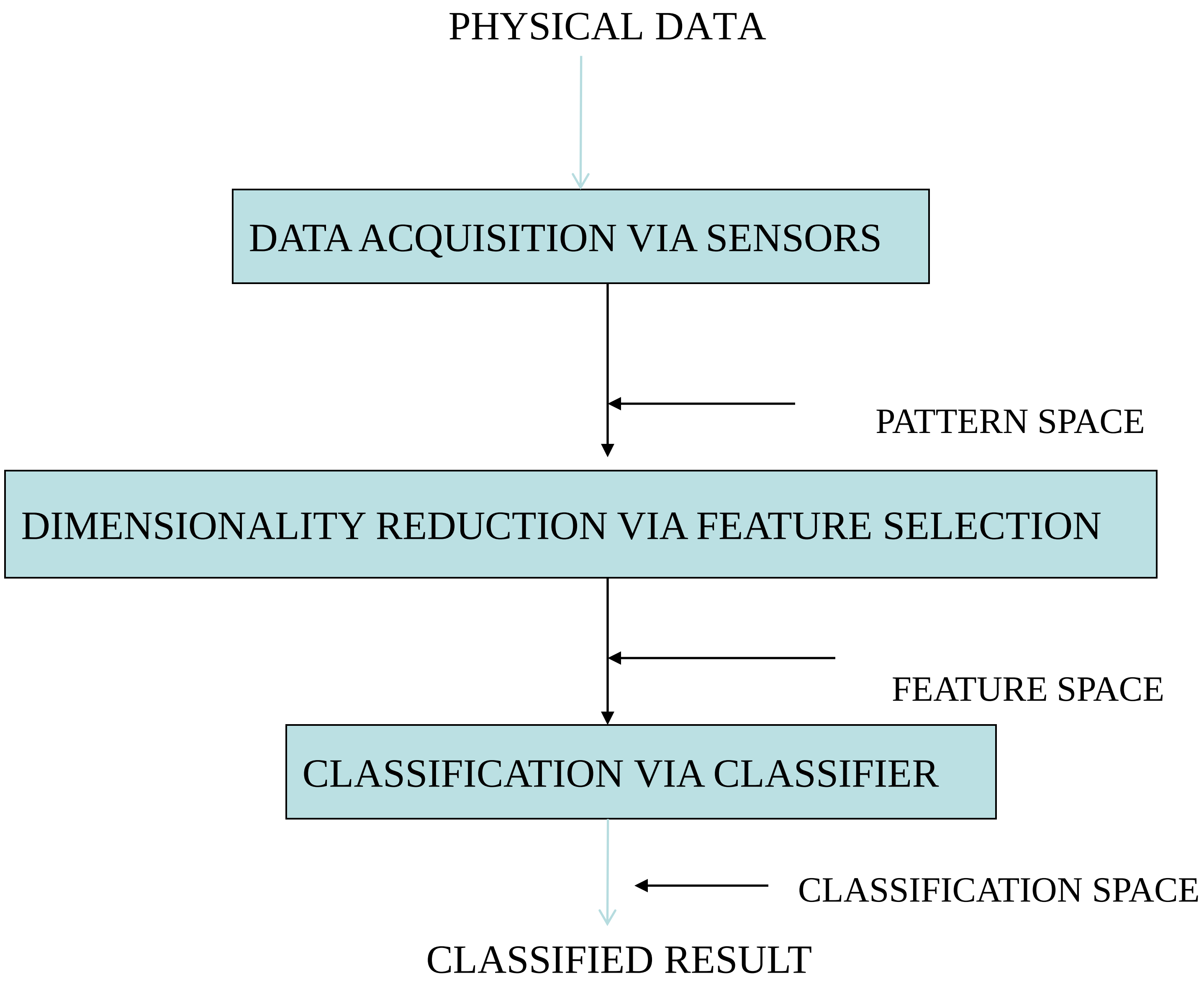
APPLICATIONS OF FUZZY THEORY

FUZZY PATTERN RECOGNITION

- In real-life we deal with complex patterns.
- Pattern recognition involves the search for structure in the complex patterns.
- Fuzzy set theory well accepted methodology for pattern recognition.
- Pattern recognition usually consists of (1) data acquisition (2) feature selection (3) classification.

UNCERTAINTY IN PATTERN RECOGNITION PROCESS

- Uncertainty in the feature space.
- Uncertainty in the classification space.
- Human beings express and process data in nonnumerical levels.
- Classification is performed by relationships.
- Classification rule: If an object is *heavy* and *small* and it moves *fast*, then it probably belong to the class ω_i


PHYSICAL DATA
DATA ACQUISITION VIA SENSORS
PATTERN SPACE
DIMENSIONALITY REDUCTION VIA FEATURE SELECTION
FEATURE SPACE
CLASSIFICATION VIA CLASSIFIER
CLASSIFICATION SPACE
CLASSIFIED RESULT

Fuzzy Pattern Recognition

- Pattern recognition partitions a set of patterns into classes depending on the similarity in features of the patterns.
- If Features are not very distinct, pattern classes overlap.
- Fuzzy pattern recognition capable of partitioning the patterns by soft boundaries.
- Patterns are classified into one or more classes with a certain degree of membership to belong to each class.
- **Fuzzy Clustering Algorithm**

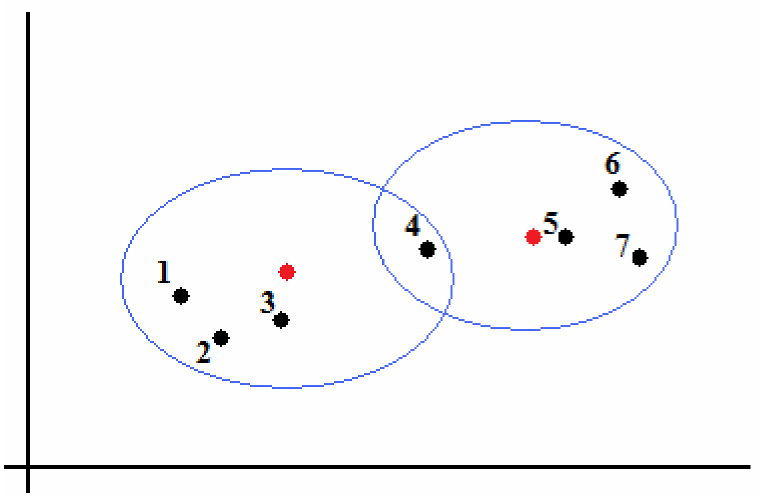
Classify a given set of p dimensional data points into a set of c fuzzy classes.

Definition

- A cluster is a group of objects that are more similar to one another than to the members of other clusters
- Similarity is often defined by means of a *distance norm.*
- Data can reveal clusters of different geometrical shapes, sizes and densities
- The performance of most clustering algorithms is influenced by the geometrical shapes, densities of the individual clusters and the spatial relations and distances among the clusters.

Clustering algorithms may use an *objective function* to measure the desirability of partitions.

- Hard Clustering vs. Soft Clustering
- Hard clustering means partition of the data into mutually exclusive subsets.
- Soft clustering allow the data points belong to different clusters.
- The discrete nature of hard clustering has some difficulties with the objective functions which are not differentiable



Fuzzy C-means Clustering

1. *Initialize* $U=[u_{ij}]$ *matrix*, $U^{(0)}$
2. *At k-step: calculate the centers vectors* $C^{(k)}=[c_j]$ *with* $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m}$$

m is fuzzyness index > 1

3. *Update* $U^{(k)}$, $U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

Usually, *m* = 2

$$\mu_{ik} \in [0,1], \quad 1 \le i \le c, \quad 1 \le k \le N,$$

$$\sum_{i=1}^{c} \mu_{ik} = 1, \quad 1 \le k \le N,$$

$$0 < \sum_{k=1}^{N} \mu_{ik} < N, \quad 1 \le i \le c.$$

4. *If* $\| U^{(k+1)} - U^{(k)} \| < \varepsilon$ *th*

2

Step 1: Initialize of clusters randomly.

Cluster	(1, 3)	(2, 5)	(4, 8)	(7, 9)
1)	0.8	0.7	0.2	0.1
2)	0.2	0.3	0.8	0.9

Step 2: Find out the centroid.

```
V11  = (0.8^2 *1 + 0.7^2 * 2 + 0.2^2 * 4 + 0.1^2 * 7) / ( (0.8^2 + 0.7^2  + 0.2^2  + 0.1^2 ) = 1.568
V12  = (0.8^2 *3 + 0.7^2 * 5 + 0.2^2 * 8 + 0.1^2 * 9) / ( (0.8^2 + 0.7^2  + 0.2^2  + 0.1^2 ) = 4.051
V21  = (0.2^2 *1 + 0.3^2 * 2 + 0.8^2 * 4 + 0.9^2 * 7) / ( (0.2^2 + 0.3^2  + 0.8^2  + 0.9^2 ) = 5.35
V22  = (0.2^2 *3 + 0.3^2 * 5 + 0.8^2 * 8 + 0.9^2 * 9) / ( (0.2^2 + 0.3^2  + 0.8^2  + 0.9^2 ) = 8.215
```

Centroids are: (1.568, 4.051) and (5.35, 8.215)

Step 3: Find out the distance of each point from the centroid.

$D11 = ((1 - 1.568)^2 + (3 - 4.051)^2)^{0.5} = 1.2$

$D12 = ((1 - 5.35)^2 + (3 - 8.215)^2)^{0.5} = 6.79$

Similarly, the distance of all other points is computed from both the centroids.

- Step 4: Updating membership values.

For point 1 new membership values are:

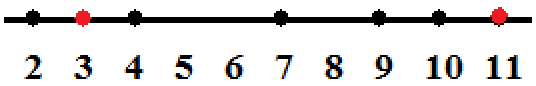
γ_{11}= [{ [$(1.2)^2$ / $(1.2)^2$] + [$(1.2)^2$ / $(6.79)^2$]} ^ {(1 / (2 – 1))}] $^{-1}$ = 0.96

γ_{12}= [{ [$(6.79)^2$ / $(6.79)^2$] + [$(6.79)^2$ / $(1.2)^2$]} ^ {(1 / (2 – 1))}] $^{-1}$ = 0.04

- **Step 5:** Repeat the steps(2-4) until the constant values are obtained for the membership values or the difference is less than the tolerance value (a small value up to which the difference in values of two consequent updations is accepted).
- **Step 6:** Defuzzify the obtained membership values.

Fuzzy C-means Clustering

- For example: we have initial centroid 3 & 11 (with m=2)



u_{ij} is the degree of membership of x_i in the cluster j.

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{\frac{2}{m-1}}}$$

- For node 2 (1st element):

$$U_{11} = \frac{1}{\left(\frac{2-3}{2-3}\right)^{\frac{2}{2-1}} + \left(\frac{2-3}{2-11}\right)^{\frac{2}{2-1}}} = \frac{1}{1+\frac{1}{81}} = \frac{81}{82} = 98.78\%$$

The membership of first node to first cluster

$$U_{12} = \frac{1}{\left(\frac{2-11}{2-3}\right)^{\frac{2}{2-1}} + \left(\frac{2-11}{2-11}\right)^{\frac{2}{2-1}}} = \frac{1}{81+1} = \frac{1}{82} = 1.22\%$$

The membership of first node to second cluster