Soft Computing

Unit - 1

Fuzzy Logic

Introduction to Neuro - Fuzzy and Soft Computing ; Fuzzy Sets - Basic Definition and Terminology - Set Theoretic Operations - Member Function Formulation and Parametrization - Fuzzy Rules and Fuzzy Reasoning - Fuzzy Inference Systems - Mamdani Fuzzy Model - Sugeno Fuzzy Models - Tsukamoto Fuzzy Models - Input Space Partitioning & Fuzzy Model

* Introduction to Neuro Fuzzy and Soft Computing

→ Soft - Computing : an approach to construct computationally intelligent systems. These intelligent systems possess human-like expertise within a specific domain, adapt themselves and learn to do better in changing environments, & explain how they make decisions or take actions.

→ Neuro - Fuzzy Computing : Neural networks that recognize patterns and adapt themselves to cope with changing environments. Fuzzy inference systems incorporate human knowledge and perform inferencing & decision making.

* Constituents of Soft Computing?

Methodology	Strength
(i) Neural network	learning & adaptation
(ii) Fuzzy set theory	knowledge representation via fuzzy if-then rules
(iii) Genetic algorithms & simulated annealing	systematic random search
(iv) Conventional AI	symbolic manipulation

* Artificial Intelligence and Expert Systems

Definitions for AI

(i) AI is the study of agents that exist in an environment and perceive and act.

(ii) the art of making computers do smart things

(iii) a programming style, where programs operate on data according to rules in order to accomplish rules.

(iv) providing computers with the ability to display behavior that would be regarded as intelligent if it were observed in humans.

Definitions of Expert Systems

→ (i) a computer program using expert knowledge to attain high levels of performance in a narrow problem area.

(ii) a caricature of a human expert

* Why Neuro-Fuzzy Models

→ The human brain interprets imprecise and incomplete sensory information.

→ Fuzzy set theory provides a systematic calculus to deal with such information, and it performs numerical computation by using labels stipulated by membership functions.

→ A selection of fuzzy if-then rules is a key component ~~of~~ a fuzzy inference system.

→ However, fuzzy inference systems lack the adaptability to deal with changing external environments.

→ Thus, neural network learning concepts are incorporated in fuzzy inference systems, resulting in neuro-fuzzy modeling.

* Evolutionary Computation

→ Evolution propels living systems' toward higher-level intelligence.

→ Evolutionary computing techniques are based on the principle of natural selection (like genetic algorithms).

→ deals with survival of the fittest among individuals of consecutive ~~operation~~ generations for solving a problem.

* Neuro - Fuzzy and Soft Computing Characteristics

① Human expertise - in the form of

(i) fuzzy if-then rules

(ii) conventional knowledge representation

② Biologically inspired computing model

- inspired by biological neural networks
- ANN used for perception, pattern recognition

③ New Optimization Techniques

- genetic algorithms
- simulated annealing
- random search method
- downhill simplex method

④ Numerical Computation

- relies mostly on numerical computation than like symbolic AI

⑤ New Application Domains

used in

- adaptive signal processing
- non linear regression
- pattern recognition

(6) Model-free learning

- have the capacity to construct models using only target system sample data.

(7) Intensive computation

- does not assume too much background knowledge
- relies heavily on high-speed number crunching computation to find rules

(8) Fault Tolerance

- deletion of a neuron in a neural network, does not necessarily destroy the system
- system continues to perform because of its parallel and redundant architecture, although performance quality gradually decreases

(9) Goal driven characteristics

- path from current state to the solution does not matter as long as one is moving to the goal in the long run.
- true in the case of genetic algorithms, simulated annealing & random-search method.

(10) Real world applications

- soft computing is an integrated approach that utilizes specific techniques within subtasks to construct satisfactory solutions.

*Fuzzy Set Theory

A. Classical Set : A set with a crisp boundary, there is a clear unambiguous boundary.
The flaw in classical sets is that there is a sharp transition between inclusion and exclusion in a set.

B. Fuzzy Set : a set without a crisp boundary.
The transition from belonging to a set and not belonging to a set is gradual, and this smooth transition is characterized by membership functions flexibility.

* Definitions

① Classical Set

A classical set A, $A \subseteq X$, is defined as a collection of elements $x \in X$, such that each x can either belong or not belong to the set A.
It can be represented by a set of ordered pairs $(x,0)$ or $(x,1)$ indicating $x \notin A$ and $x \in A$ respectively.

② ~~Fuzzy Set~~

* Fuzzy Sets and Membership Functions

If X is a collection of objects denoted by x, then a fuzzy set A in X is defined as:

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

$\mu_A(x)$ is called the membership function (MF). The MF maps each element of X to a membership grade/value between 0 and 1.

* Fuzzy sets in a discrete non-ordered universe

The elements are distinct and separate, but there is no inherent order between them.

eg. X = { San Francisco, Boston, LA }

C = "desirable city to live in"

C = { (San Francisco, 0.9), (Boston, 0.8), (LA, 0.6) }

* Fuzzy sets in a discrete ordered universe

Unlike in a non-ordered universe, the order of elements in the set is significant.

For eg. Let $X = \{1, 2, 3, 4, 5, 6\}$ be the no. of children a family may choose to have.

The fuzzy set A is "sensible no. of children in a family"

then $A = \{(0, 0.1), (1, 0.3), (2, 0.7), (3, 1), (4, 0.7), (5, 0.3), (6, 0.1)\}$

* Fuzzy Sets with a Continuous Universe

→ where the membership function can be expressed as a function corresponding to continuous values.

Example: Let R^+ be the set of possible ages for human beings.

The Fuzzy set B is "about 50 yrs old"

$$B = \{ (x, \mu_B(x)) \mid x \in X \}$$

where $\mu_B(x) = \dfrac{1}{1 + \left(\dfrac{x-50}{10}\right)^4}$

* Difference between fuzzy sets and probability theory?

→ In fuzzy sets, the specification of membership functions is subjective, different individuals may perceive it differently.

→ Probability on the other hand deals with the objective treatment of random phenomena

∴ The subjectivity and non-randomness of fuzzy sets is the primary difference.

* Representation of discrete and continuous fuzzy sets

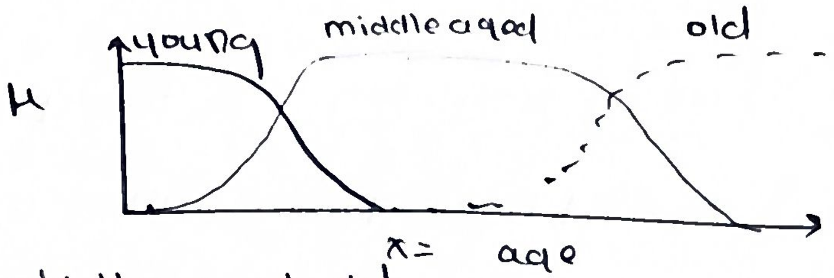
$$A = \begin{cases} \sum_{x_i \in X} \mu_A(x_i) / x_i & \rightarrow \text{discrete} \\ \int_x \mu_A(x) / x & \rightarrow \text{continuous} \end{cases}$$

Σ, $\int$ stand for the union of $(x, \mu_A(x))$ – they do not indicate summation or integration

* Linguistic Values

→ Fuzzy sets that carry names that conform to adjectives appearing in our daily linguistic usage, (say like - small, medium, large) are called __linguistic values__ or __linguistic labels__.

→ Thus, the universe of discourse X is often called the __linguistic variable__.



eg. Linguistic variable = Age

Linguistic values = young, middle aged, old

* Nomenclature

① __Support__ - The support of a fuzzy set is the set of all points x in X such that $\mu_A(x) > 0$.

$$\text{support}(A) = \{ x \mid \mu_A(x) > 0 \}$$

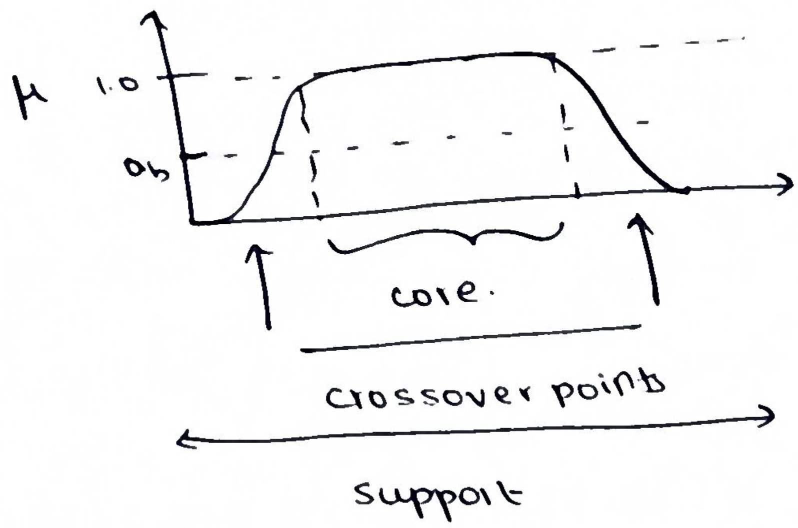
② __Core__ : The core of a fuzzy set A is the set of all points x in X such that $\mu_A(x) = 1$

$$\text{core}(A) = \{ x \mid \mu_A(x) = 1 \}$$

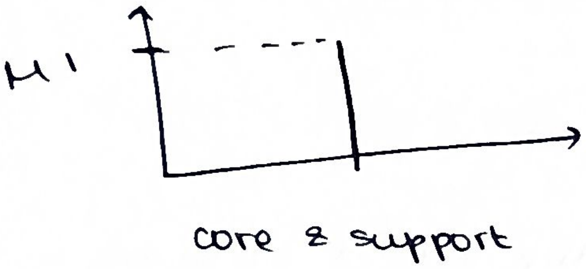
③ __Normality__ : A fuzzy set A is normal if its core is non-empty. i.e we can always find a point $x \in X$ such that $\mu_A(x) = 1$

④ __Crossover points__: A crossover point of a fuzzy set A is a point $x \in X$ at which $\mu_A(x) = 0.5$

$$\text{crossover}(A) = \{ x \mid \mu_A(x) = 0.5 \}$$



(5) <u>Fuzzy singleton</u> : A fuzzy set whose support is a single point in X with $\mu_A(x) = 1$ is called a fuzzy singleton



* Cuts

(1) <u>α-cut, α-level set</u> - The α-cut or α-level set of a fuzzy set A is a crisp set defined by

$$A_\alpha = \{x \mid \mu_A(x) \geq \alpha\}$$

(2) <u>Strong α-cut or strong α-level set</u> - It is defined as

$$A'_\alpha = \{x \mid \mu_A(x) \boxed{\geq} \alpha\}$$ (note the equality sign)

* Convexity of Fuzzy Sets

→ Convexity describes how spread out or how smooth a membership function of a fuzzy set is.

→ Consider the example of a glass of water & a fuzzy set showing the wetness of objects. The convex fuzzy set would be like smoothly pouring water over a range of objects. On the other hand, a concave fuzzy set would be like randomly splashing water on objects. Some may be wet, some may be dry, and there is no clear pattern of wetness in between.

Mathematically, a fuzzy set A is convex iff for any $x_1, x_2 \in X$ and $\lambda \in [0,1]$

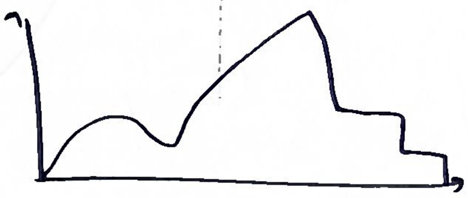
$$\mu_A(\lambda x_1 + (1-\lambda) x_2) \geq \min\{\mu_A(x_1), \mu_A(x_2)\}$$

→ $\mu_A(x_1)$ and $\mu_A(x_2)$ are the membership values of the fuzzy set A for elements x_1 and x_2.

→ λ is a weight that determines the degree of convexity. It represents how much one mixes or interpolates between the membership values of x and y.

Example

convex



non-convex

* Fuzzy Numbers and Bandwidth

→ A fuzzy number A is a fuzzy set in the real line that satisfies the conditions for normality and convexity

normality ↓ at least $\mu_A(x) = 1$

* Bandwidths of normal and convex fuzzy sets

→ For a normal and convex fuzzy set, the bandwidth or width is defined as the distance between the two unique crossover points

$$width(A) = |x_2 - x_1|$$

where $\mu_A(x_1) = \mu_A(x_2) = 0.5$

* Symmetry

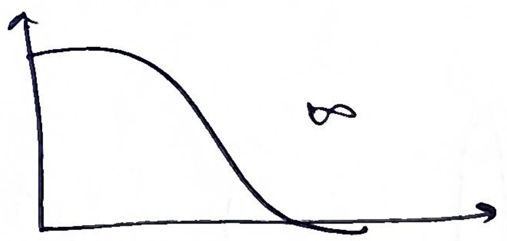
A fuzzy set A is symmetric if its MF is symmetric around a certain point, say at $x = c$

i.e $\mu_A(x+c) = \mu_A(x-c) \quad \forall x \in X$

* Open left, open right, closed

open left : $\lim_{x \to -\infty} \mu_A(x) = 1$

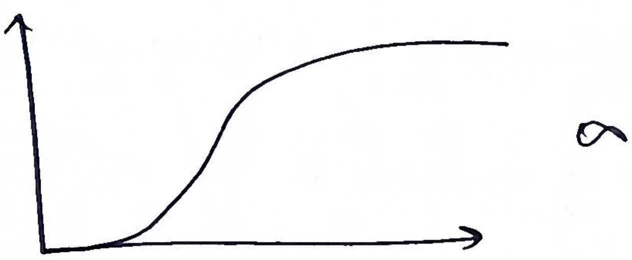
and $\lim_{x \to +\infty} \mu_A(x) = 0$



open right

$\lim_{x \to -\infty} \mu_A(x) = 0$

$\lim_{x \to +\infty} \mu_A(x) = 1$

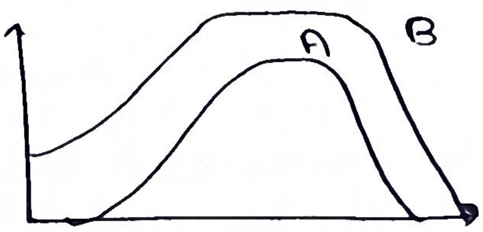


closed $\lim_{x \to -\infty} \mu_A(x) = \lim_{x \to +\infty} \mu_A(x) = 0$

* Set Theoretic Operations

① Containment or Subset : Fuzzy set A is contained in fuzzy set B iff $\mu_A(x) \le \mu_B(x)$

$$A \subseteq B \iff \mu_A(x) \le \mu_B(x)$$



② Union or Disjunction : The union of 2 fuzzy sets A and B is a fuzzy set C is $C = A \cup B$, whose MF is related by:

$$\mu_C(x) = \max(\mu_A(x), \mu_B(x)) = \mu_A(x) \vee \mu_B(x)$$

→ Smallest fuzzy set contained in both A & B

③ Intersection or Conjunction : The intersection of 2 fuzzy sets A and B is $C = A \cap B$, whose MF is related by:

$$\mu_C(x) = \min(\mu_A(x), \mu_B(x)) = \mu_A(x) \wedge \mu_B(x)$$

→ Largest fuzzy set contained in both A & B.

④ Negation or Complement

The complement of fuzzy set A, denoted by $\bar{A}$ is donoted by

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$$

* Cartesian Product and Co Product : If A & B are fuzzy sets in A and B.

① Cartesian product

The cartesian product of A and B denoted by $A \times B$ is a fuzzy set with the membership function

$$\mu_{A \times B}(x, y) = \min(\mu_A(x), \mu_B(y))$$

(2) Cartesian co-product : The cartesian co product is a fuzzy set with the membership function:

$$\mu_{A+B}(x,y) = \max(\mu_A(x), \mu_B(y))$$

* MF Formulation and Parametrization

→ It is convenient to define a membership function (MF) by expressing it as a mathematical function

→ The derivatives of the MF can be defined with respect to their inputs and parameters.

→ These parameters are used for fine tuning fuzzy inference systems to achieve a desired input/output mapping.

(1) MFs of One Dimension.

A. Trianqular MFs

A trianqular MF is specified by 3 parameters {a, b, c} as follows:

$$\text{triangle}(x; a, b, c) = \begin{cases} 0 & x \leq a. \\ \frac{x-a}{b-a} & a \leq x \leq b \\ \frac{c-x}{c-b} & b \leq x \leq c \\ 0 & c \leq x. \end{cases}$$

An alternative expression using min & max is

$$\text{triangle}(x; a, b, c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$$

{a, b, c} determine the x coordinates of the 3 corners of the underlying triangle

(2) Trapezoidal MFs

→ A trapezoidal MF is specified by 4 parameters $\{a, b, c, d\}$

$$\text{trapezoid}(x; a, b, c, d) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x \le c \\ \frac{d-x}{d-c} & c \le x \le d \\ 0 & d \le x. \end{cases}$$

using min and max, the expression would be:

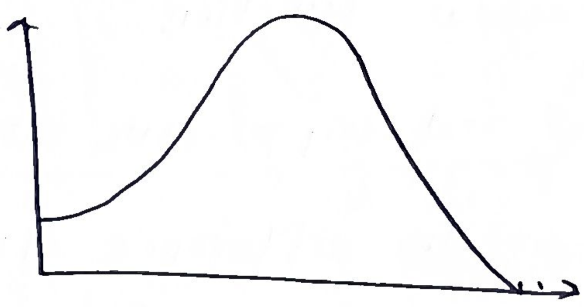
$$\text{trapezoid}(x; a, b, c, d) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right)$$

The parameters $\{a, b, c, d\}$ determine the x coordinates of the 4 corners of the trapezoidal MF.

(3) Gaussian MFs

A gaussian MF is specified by 2 parameters $\{c, \sigma\}$

$$\text{gaussian}(x; c, \sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2}$$



c → represents MFs center

σ → determines MFs width

(4) Generalized Bell MFs

A bell MF is specified by 3 parameters $\{a, b, c\}$

$$\text{bell}(x; a, b, c) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}}$$

→ b is usually +ve - if -ve, it becomes an upside-down bell

→ is a direct generalization of the Cauchy distribution - is called the Cauchy MF.

(5) Sigmoidal MF

→ A sigmoidal MF is defined by:

$$sig(x; a, c) = \frac{1}{1 + \exp[-a(x-c)]}$$

where a controls the slope at the crossover point $x=c$.

→ Depending on the sign of the parameter a, a sigmoidal function is inherently open left or right, and is thus used to represent concepts such as "very large" or "very negative".

Usage

→ used as the activation function in many artificial neural networks

→ For a neural network to simulate the behavior of a fuzzy inference - one must synthesize a closed MF through a sigmoidal function.

Closed and asymmetric MFs based on sigmoidal functions

(i) take the difference $|y_1 - y_2|$

(ii) take the product $y_1 * y_2$

(6) Left-right MF (L-R MF)

An L-R MF is specified by 3 parameters $\{\alpha, \beta, c\}$

$$LR(x; c, \alpha, \beta) = \begin{cases} F_L\left(\frac{c-x}{\alpha}\right), & x \le c \\ F_R\left(\frac{x-c}{\beta}\right), & x \ge c. \end{cases}$$

where $F_L(x)$ and $F_R(x)$ are monotonically decreasing functions defined on $[0, \infty)$.

with $F_L(0) = F_R(0) = 1$

$$\lim_{x \to \infty} F_R(x) = \lim_{x \to \infty} F_R(x) = 0.$$

Note: The L-R MF is extremely flexible in specifying fuzzy set. However, it is not often used in practice, because of its unnecessary complexity.

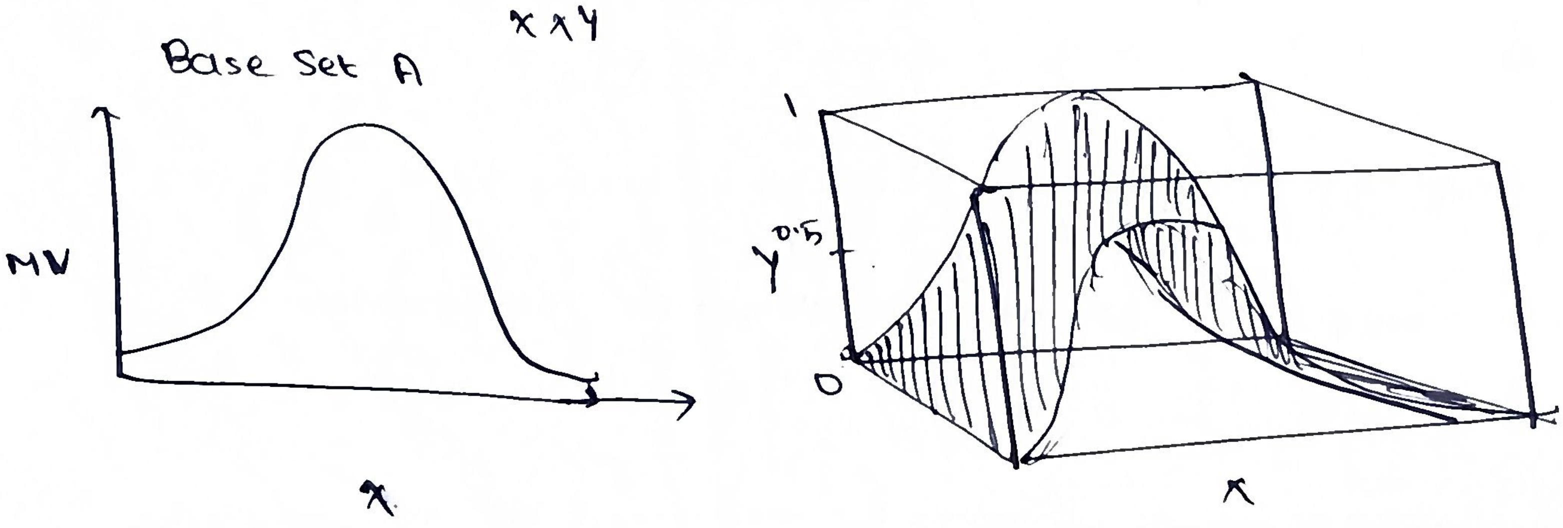
* MFs of Two Dimensions

→ It is advantageous to use MFs with two inputs, each in a different universe of discourse, called 2-D MFs.

→ One natural way to extend 1D MFs to 2D MFs is via cylindrical extension.

→ If A is a fuzzy set in X, then its cylindrical extension in X × Y is a fuzzy set c(A) defined by

$$c(A) = \int_{X \times Y} \mu_A(x) / (x, y)$$



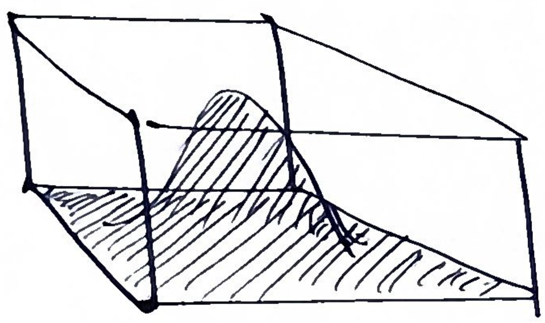
Projections of Fuzzy Sets

→ The operation of projection decreases the dimensions of a given multidimensional membership function.

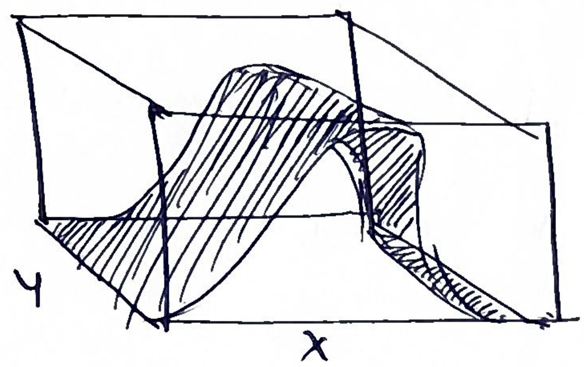
→ Let R be a two-dimensional fuzzy set on $X \times Y$. Then the projections of R onto X and Y are defined as:

$$R_X = \int_X \left[\max \mu_R(x,y)\right] / x$$

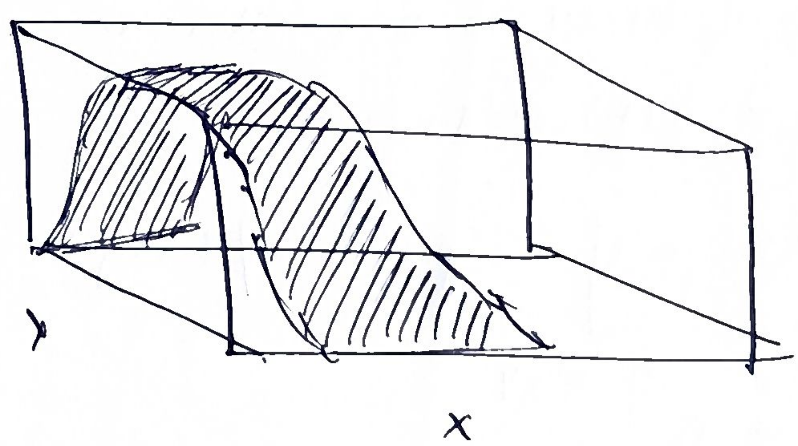
$$R_Y = \int_Y \left[\max \mu_R(x,y)\right] / y$$



2D MF



Projection onto X



Projection onto Y

* Composite and Non-Composite MFs

→ MFs of 2 dimensions fall into 2 categories: composite and non-composite

→ If an MF of two dimensions can be expressed as an analytic expression of two MFs of one dimension, then it is composite, otherwise it is non- composite.

→ A composite 2D MF is usually the result of 2 statements joined by the AND or OR connectives.

eg. If fuzzy set A = (x,y) is near (3,4)

$$\mu_A(x,y) = \exp\left[-\left(\frac{x-3}{2}\right)^2 - (y-4)^2\right]$$

$$= \exp\left[-\left(\frac{x-3}{2}\right)^2\right] \exp\left[-\left(\frac{y-4}{1}\right)^2\right]$$

$$= \text{gaussian}(x; 3,2)\ \text{gaussian}(y; 4,1)$$

= composite

gaussian

$$e^{-1/2\left(\frac{x-c}{\sigma}\right)^2}$$

$$= (x; c, \sigma)$$

but,

$$\mu_A(x,y) = \frac{1}{1+|x-3||y-4|^{2.5}}$$ is non-composite.

* Composite 2D MFs based on min and max operators

→ When the min operator is used to aggregate one - dimensional MFs, the resulting 2D MFs can be viewed either as a result of applying classical fuzzy intersection to the cylindrical extensions of each 1D MF, or as a Cartesian product of two 1D fuzzy sets

→ The same interpretation applies to the max operator as well.

* Derivatives of Parameterized MFs

→ To make a fuzzy system adaptive, we need to know the derivatives of an MF w.r.t to its input parameters.

Example 1

$$y = \text{gaussian}(x ; \sigma, c) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2}$$

$$\frac{\partial y}{\partial x} = -y \cdot \frac{\partial y}{\partial x}\left(\frac{(x-c)^2}{\sigma^2}\right)$$

$$= y \cdot \frac{-2(x-c)}{2\sigma}$$

$$= \frac{-(x-c) \cdot y}{\sigma}$$

$$\frac{\partial y}{\partial \sigma} = y \cdot \frac{\partial}{\partial \sigma}\left(\frac{(x-c)^2}{\sigma^2}\right)$$

$$= y \cdot \frac{(x-c)^2}{\sigma^3}$$

$$\frac{\partial y}{\partial c} = \frac{x-c}{\sigma^2} \cdot y$$

Example 2 $\quad y = \text{bell}(x : a, b, c) = \dfrac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}}$

$$\frac{\partial y}{\partial x} = \begin{cases} \dfrac{-2b}{x-c} \cdot y(1-y) & \text{if } x \neq c \\ 0 & \text{if } x = 0 \end{cases}$$

do for $\frac{\partial y}{\partial a}$, $\frac{\partial y}{\partial b}$, $\frac{\partial y}{\partial c}$

* Extension Principle

→ The extension principle provides a general procedure for extending crisp domains of mathematical expressions to fuzzy domains.

→ This procedure generalizes a common point-to-point mapping of a function $f(\cdot)$ to a mapping between fuzzy sets

→ If f is a function from X to Y, and X is a fuzzy set on X defined as

$$A = \mu A(x_1)/x_1 + \mu A(x_2)/x_2 + \cdots \mu A(x_n)/x_n$$

Then, the image of fuzzy set A under the mapping $f(\cdot)$ is

$$B = f(A) = \mu A(x_1)/y_1 + \mu A(x_2)/y_2 + \cdots \mu A(x_n)/y_n$$

→ If there are multiple points mapped, then take the max value.

i.e The membership grade of B (at $y = y^*$) is the maximum of the membership grades of A at $x = x_1$, $x = x_2$ .. etc.

$$\text{i.e } \mu B(y) = \max_{x = f^{-1}(y)} \mu A(x)$$

Example:- Apply the extension principle to the given fuzzy set

$$A = 0.1/-2 + 0.4/-1 + 0.8/0 + 0.9/1 + 0.3/2$$

$$f(x) = x^2 - 3$$

Ans

$f(0) = -3$

$f(-2) = 1$

$f(-1) = -2$

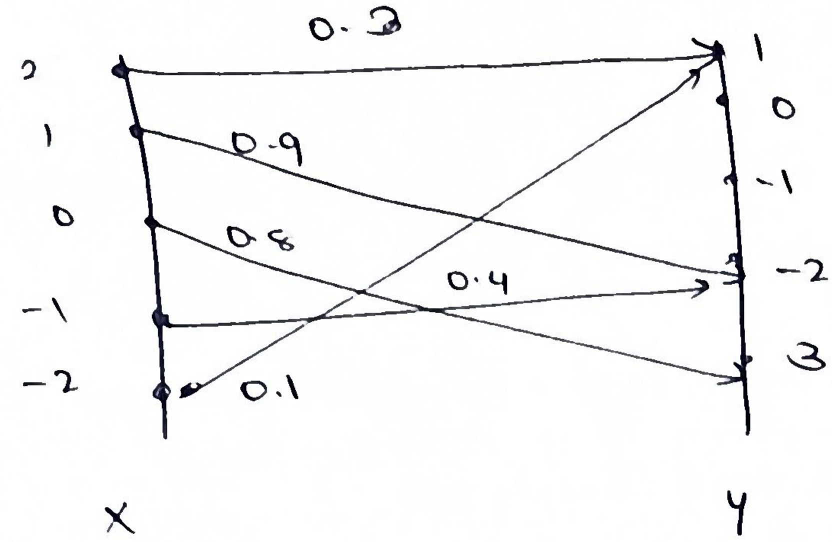
$f(1) = -2$

$f(2) = 1$

$$B = 0.1/1 + 0.4/-2 + 0.8/-3 + 0.9/-2 + 0.3/1$$

~~B = 0.1/1~~ $B = (0.1 \vee 0.3)/1 + (0.4 \vee 0.9)/-2 + 0.8/-3$

$$\boxed{B = 0.8/-3 + 0.3/1 + 0.9/-2}$$



Alt. Definition

Suppose that the function f is a mapping from an n-dimensional Cartesian product space $X_1 \times X_2 \times \dots X_n$ to a one-dimensional universe Y such $y = f(x_1, \dots x_n)$, and if $A_1 \dots A_n$ are fuzzy sets in $X_1 \dots X_n$.

Then $\mu_B(y) = \begin{cases} \max\left[\min \mu_{A_i}(x_i)\right] & \text{if } f^{-1}(y) \neq 0 \\ 0 & \text{if } f^{-1}(y) = 0 \end{cases}$

* Fuzzy Relations

→ Fuzzy relations are fuzzy sets which map each element in $X \times Y$ to a membership grade between 0 & 1.

→ Unary fuzzy relations have 1D membership functions

Binary fuzzy relations have 2D membership functions.

→ For a binary fuzzy relation, if X & Y are 2 universes of discourse

$R = \{((x,y), \mu_R(x,y)) \mid (x,y) \in X \times Y\}$

$\mu_R(x,y) = 2D$ MF

Example of a binary fuzzy relation

Let $X = Y = R^+$ and R = "y is much greater than x". The MF is defined as:

$$\mu_R(x,y) = \begin{cases} \dfrac{y-x}{x+y+2} & \text{if } y > x \\ 0 & \text{if } y \le x \end{cases}$$

If $x = \{3, 4, 5\}$ and $y = \{3, 4, 5, 6, 7\}$

The relation matrix would be:

$$R = \begin{bmatrix} 0 & 0.111 & 0.200 & 0.273 & 0.333 \\ 0 & 0 & 0.041 & 0.167 & 0.231 \\ 0 & 0 & 0 & 0.077 & 0.143 \end{bmatrix}$$

Example2 x is close to y (numbers)

x depends on y (events)

x & y look alike (persons, objects)

If x is large, then y is small (x is an observed reading, and y is a corresponding action)

↓

used in fuzzy inference systems

* Operations on Fuzzy Relations

① Max-Min Composition

Let R_1 and R_2 be two fuzzy relations defined on $X \times Y$ and $Y \times Z$. The max-min composition of R_1 and R_2 is a fuzzy set defined by:

$$R_1 o R_2 = \{ [(x,z), \max_y \min(\mu_{R_1}(x,y), \mu_{R_2}(y,z))] | x \in X, y \in Y, z \in Z \}$$

(OR)

$$\mu_{R_1 o R_2}(x,z) = \max_y \min [\mu_{R_1}(x,y), \mu_{R_2}(y,z)]$$

$$= \vee_y [\mu_{R_1}(x,y) \wedge \mu_{R_2}(y,z)]$$

$\vee = \max$, $\wedge = \min$

→ also called the max-min product

(2) Max-product composition

The max-product composition is defined as follows:

$$\mu_{R_1 o R_2}(x,z) = \max_y [\mu_{R_1}(x,y) \, \mu_{R_2}(y,z)]$$

* Properties of binary relations

(i) Associativity : $Ro(SoT) = (RoS)oT$

(ii) Distributivity over union : $Ro(S \cup T) = (RoS) \cup (RoT)$

(iii) Weak distributivity over intersection : $Ro(S \cap T) \subseteq (RoS) \cap (RoT)$

(iv) Montonicity : $S \subseteq T \Rightarrow RoS \subseteq RoT$

Example1 Let $R_1 = x$ is relevant to y
$R_2 = y$ is relevant to z be a fuzzy relations on $X \times Y$ and $Y \times Z$, where $X = \{1, 2, 3\}$, $Y = \{\alpha, \beta, \gamma, \delta\}$ and $Z = \{a, b\}$

R_1 and R_2 are expressed by the 2 relation matrices:

$$R_1 = \begin{array}{c} \\ 2 \\ \\ \end{array}\begin{bmatrix} 0.1 & 0.3 & 0.5 & 0.7 \\ 0.4 & 0.2 & 0.8 & 0.9 \\ 0.6 & 0.8 & 0.3 & 0.2 \end{bmatrix} \qquad R_2 = \begin{array}{c} \begin{array}{cc} a & \end{array} \\ \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.3 \\ 0.5 & 0.6 \\ 0.7 & 0.2 \end{bmatrix} \end{array}$$

Find the degree of relevance between 2 and a.

Ans: using max-min composition

$$\mu_{R_1 \circ R_2}(2, a) = \max\{\min\{(0.4, 0.9), (0.2, 0.2), (0.8, 0.5), (0.9, 0.7)\}\}$$

$$= \max\{0.4, 0.2, 0.5, 0.7\}$$

$$= \underline{\underline{0.7}}$$

using max-product composition

$$\mu_{R_1 \circ R_2}(2, a) = \max\{(0.4 \times 0.9), (0.2 \times 0.2), (0.8 \times 0.5), (0.9 \times 0.7)\}$$

$$= \max\{0.36, 0.04, 0.40, 0.63\}$$

$$= \underline{\underline{0.63}}$$

Fuzzy If-Then Rules & Fuzzy Reasoning

* **Principle of Incompatibility** : As the complexity of a system increases, our ability to make precise and significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance become almost mutually exclusive characteristics.

* **Linguistic Variables**

→ To model human thinking, to summarize information and express it in terms of fuzzy sets.

→ A linguistic variable is characterized by a quintuple $(x, T(x), X, G, M)$

x = name of the variable

$T(x)$ = term set of x, the set of linguistic values or linguistic terms

X = universe of discourse

G = syntactic rule that generates the terms in $T(x)$

M = semantic rule which associates with each linguistic value A its meaning $M(A)$, where $M(A)$ denotes a fuzzy set in X.

Example If age is a linguistic variable, the term set T(age) could be:

$T(age) = \{$ young, not young, middle age, not old, more or less old ... $\}$

The universe of discourse, X = a fuzzy set w/ values ranging from [0,100]

* Linguistic Terminology

- primary terms = base words = young, middle aged, old
- negation = not = not old $\quad NOT(A) = \neg A = \int_X [1 - \mu_A(x)] / x$
- hedges = very, more or less, quite, extremely
- connectives = and, or, either, neither. $\quad$: $A \text{ AND } B = A \cap B = \int_X [\mu_A(x) \wedge \mu_B(x)] / x$

$A \text{ OR } B = A \cup B = \int_X [\mu_A(x) \vee \mu_B(x)] / x$

* Concentration and Dilation of Linguistic Values

Let A be a linguistic value with membership function $\mu_A(\cdot)$. Then A^k is a modified version of the original linguistic value expressed as

$$A^k = \int_X [\mu_A(x)]^k / x$$

Concentration : $CON(A) = A^2$

Dilation : $DIL(A) = A^{0.5}$

→ Con(A) and DIL(A)
↓ from applying hedge very
↓ from applying the hedge more or less

* Constructing MFs for composite linguistic

Let the membership functions for young and old be defined as follows:

$$\mu_{young}(x) = bell(x, 20, 2, 0) = \frac{1}{1+\left(\frac{x}{20}\right)^4}$$

$$\mu_{old}(x) = bell(x, 30, 3, 100) = \frac{1}{1+\left(\frac{x-100}{30}\right)^6}$$

A. more or less old $= DIL(\overset{old}{A}) = \int_x \sqrt{\frac{1}{1+\left(\frac{x-100}{30}\right)^6}} \Big/ x$

B. not young and not old $= \neg young \cap \neg old$

$$= \int_x \left[1 - \frac{1}{1+\left(\frac{x}{20}\right)^4}\right] \wedge \left[1 - \frac{1}{1+\left(\frac{x-100}{30}\right)^6}\right] \Big/ x$$

C. young but not too young

too young = very young = CON(young)

$= young^2$

not too young $= \neg young^2$

$= young \wedge \neg young^2$

$$= \int_x \left[\frac{1}{1+\left(\frac{x}{20}\right)^4}\right] \wedge \left[1 - \left(\frac{1}{1+\left(\frac{x}{20}\right)^4}\right)^2\right] \Big/ x$$

D. extremely old

$$CON(CON(CON(old))) = (old)^8 = \int_x \left[\frac{1}{1+\left(\frac{x-100}{30}\right)^6}\right]^8 \Big/ x$$

* Contrast Intensification

→ The operation of contrast intensification on a linguistic value A is defined by:

$$INT(A) = \begin{cases} 2A^2 & \text{for } 0 \le \mu_A(x) \le 0.5 \\ \neg 2(\neg A)^2 & \text{for } 0.5 \le \mu_A(x) \le 1 \end{cases}$$

→ The inverse operator of contrast intensifier is the contrast diminisher DIM.

→ The repeated application of INT reduces the fuzziness of a fuzzy set, in the extreme case, the fuzzy set becomes a crisp set with boundaries at the crossover points.

* Orthogonality

A term set $T = t_1, \ldots t_n$ of a linguistic variable x on the universe X is orthogonal if it fulfils the following property:

$$\sum_{i=1}^{n} \mu_{t_i}(x) = 1, \quad \forall x \in X$$

t_i s are convex & normal fuzzy sets defined on X.

* Fuzzy - If - Then Rules

→ A fuzzy if - then rule also known as a fuzzy rule, fuzzy implication or fuzzy conditional statement, assumes the form:

if x is A, then y is B

A, B → linguistic values defined by fuzzy sets on the universe of discourse X and Y respectively.

x is A → antecedent or premise

y is B → consequence or conclusion

eg. If pressure is high, then volume is small.

Rules

$$R = A \to B = A \times B = \int_{X \times Y} \mu_A(x) \overset{\text{T-norm}}{\tilde{*}} \mu_B(y) \,/\, (x,y)$$

material implication : $R = A \to B = \neg A \cup B$

propositional calculus : $R = A \to B = \neg A \cup (A \cap B)$

extended propositional calculus : $R = A \to B = (\neg A \cap \neg B) \cup B$

* Fuzzy Reasoning

→ Fuzzy reasoning, also known as approximate reasoning, is an inference procedure that derives conclusions from a set of fuzzy If-then rules and known facts.

* Compositional Rule of Inference

* Fuzzy Reasoning

→ The basic rule of inference in traditional logic is modus ponens according to which we can infer the truth of a proposition B from the truth of A and the implication $A \to B$

i.e $A,\ A \to B = B$

→ premise1 (fact) : x is A'
premise2 (rule) : If A is A, then y is B

consequence (conclusion) : y is B

→ However, modus ponens can also be employed in an approximate manner. for eg. If the tomato is more or less red, then the tomato is more or less ripe.

Premise1 (fact) : x is A'
premise2 (rule) : If x is A, then y is B

consequence (conclusion) : y is B'

when A' is close to A and B' is close to B.

This kind of inference procedure is called approximate reasoning or fuzzy reasoning or also as generalized modus ponens (GMP)

→ Approximate Reasoning : Formally, the statements: x is A'
If x is A then y is B is defined by

$$\mu_{B'}(y) = \max_x \min[\mu_{A'}(x),\ \mu_R(x,y)]$$

ie $A' \circ (A \to B) = B'$

Single Rule with Single Antecedent

premise 1 (fact): x is A'

premise 2 (rule): if x is A then y is B

consequence y is B'

$$\mu_{B'}(y) = [\vee_x (\mu_{A'}(x) \wedge \mu_A(x)] \wedge \mu_B(y))$$

$$= w \wedge \mu_B(y))$$

Single Rule with Multiple Antecedents

premise 1 (fact): x is A' and y is B'

premise 2 (rule) : if x is A and y is B, then z is C.

consequence : z is C'

$$C' = (A' \times B') \circ (A \times B \rightarrow C)$$

$$\mu_{C'}(z) = \vee_{x,y} [\mu_{A'}(x) \wedge \mu_{B'}(y)] \wedge (\mu_A(x) \wedge \mu_B(y) \wedge \mu_C(z)]$$

$$= \vee_{x,y} [\mu_{A'}(x) \wedge \mu_{B'}(y) \wedge \mu_A(x) \wedge \mu_B(y)] \wedge \mu_C(z))$$

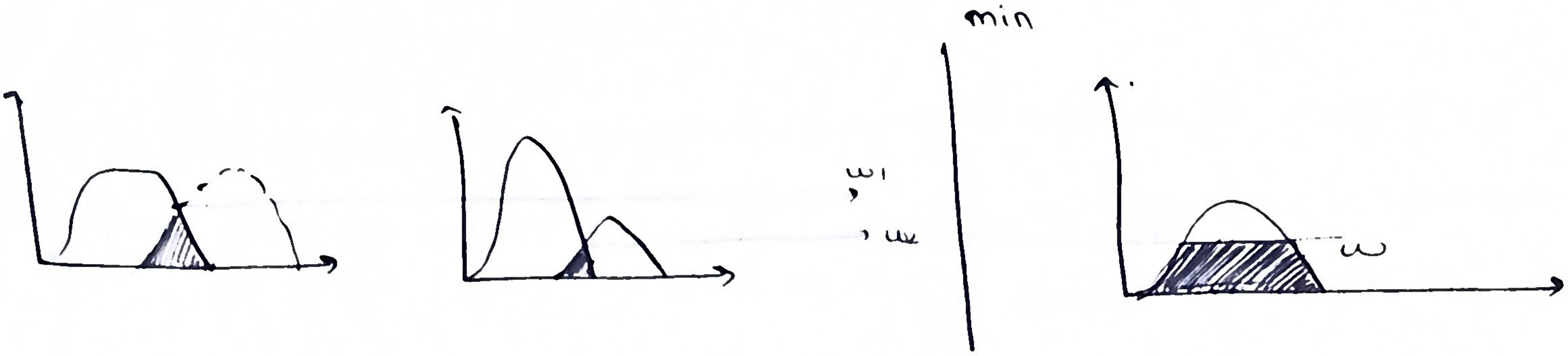
$$= \underbrace{\vee_x [\mu_{A'}(x) \wedge \mu_A(x)]}_{w_1} \wedge \underbrace{\vee_y [\mu_{B'}(y) \wedge \mu_B(y)\}}_{w_2} \wedge \mu_C(z)$$

$$= \underbrace{(w_1 \wedge w_2)}_{\text{firing strength}} \wedge \mu_C(z)$$

→ w_1 and w_2 are the maxima of the MFs of A & A' and B and B'

→ w_i denotes the degree of compatibility between A & A'

→ $w_1 \wedge w_2$ is called the firing strength or degree of fulfillment of the fuzzy rule, which represents the degree to which the antecedent part of the rule is fulfilled.



Multiple Rules with Multiple Antecedents

premise 1 (fact) : x is A' and y is B'

premise 2 (rule 1) : If x is A_1 and y is B_1 then z is C_1

premise 3 (rule 2) : If x is A_2 and y is B_2, then z is C_2

consequence (concluision) : z is C'

Verification : Let $R_1 = A_1 \times B_1 \rightarrow C_1$ and

$$R_2 = A_2 \times B_2 \longrightarrow C_2$$

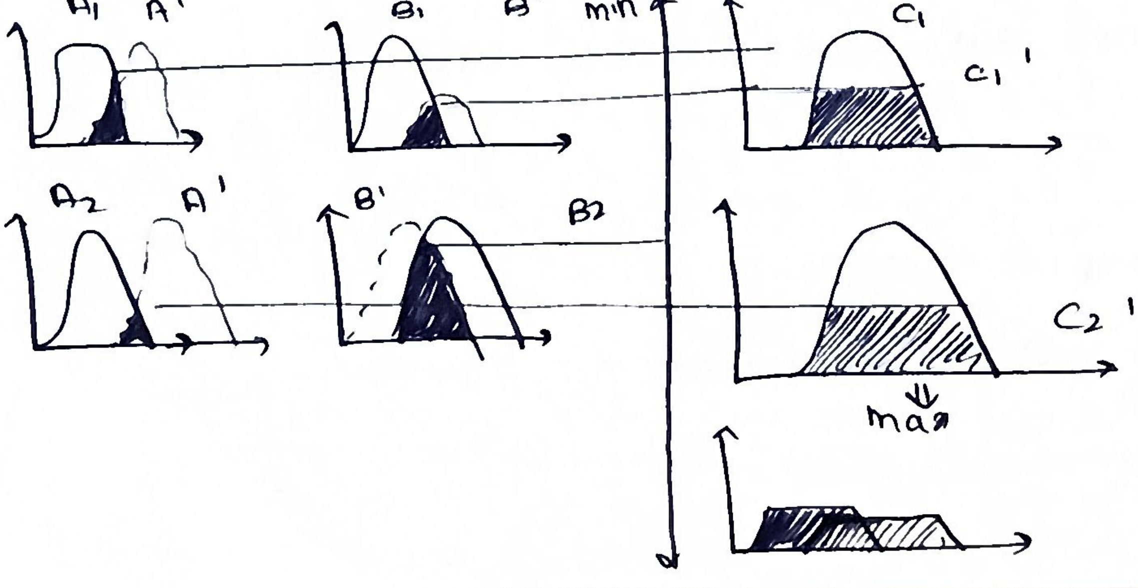
Since the max-min composition is distributive over $\cup$

$$C' = (A' \times B') \circ (R_1 \cup R_2)$$

$$= \{(A' \times B') \circ R_1\} \cup \{(A' \times B') \circ R_2\}$$

$$= C_1' \cup C_2'$$

where C_1' and C_2' are the inferred fuzzy sets for R_1 & R_2



Steps in the process of Fuzzy Reasoning

(i) Degrees of Compatibility : Compare the known facts with the antecedents of fuzzy rules to find the degrees of compatibility w.r.t each antecedent MF.

(ii) Firing Strength : Combine degrees of compatibility w.r.t antecedent MFs in a rule using fuzzy AND or OR operators to form a firing strength that indicates the degree to which the antecedent part of the rule is satisfied.

(iii) Qualified | Induced Consequent MFs : Apply the firing strength to the consequent MF to generate a qualified consequent MF. The qualified consequent MFs represent how the firing strength gets propagated and used in a fuzzy implication statement.

(iv) Overall output MF : Aggregate all of the qualified consequent MFs to obtain an overall output MF.

* Fuzzy Inference Systems

→ A computing framework based on the concepts of fuzzy set theory, fuzzy if-then rules and fuzzy reasoning

→ used in automatic control, data classification, time series prediction

→ also called fuzzy expert system
fuzzy rule based system
fuzzy model etc.

* Conceptual Components of a fuzzy inference system

(i) a rule base - contains a selection of fuzzy rules

(ii) a database or dictionary - defines the membership functions used in fuzzy rules

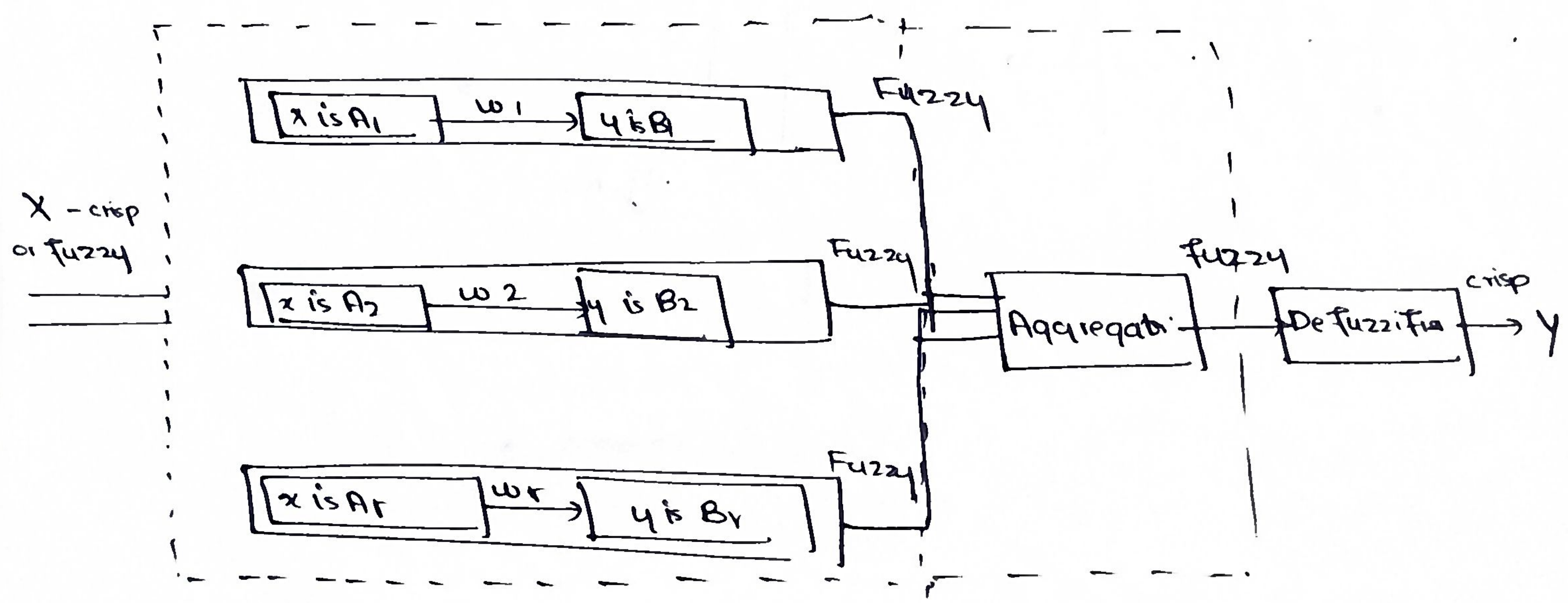
(iii) a reasoning mechanism - performs the inference procedure on the rules and given facts to derive a reasonable output or conclusion.

* Inputs and Outputs of Fuzzy Inference Systems

→ can take either fuzzy inputs or crisp inputs, but the outputs are almost always fuzzy sets

→ With crisp inputs and outputs, the fuzzy inference system implements a nonlinear mapping from its input space to output space.

→ The method of defuzzification helps ~~a crisp~~ to extract a crisp value that best represents a fuzzy set.



* Fuzzy Inference Models

(i) Mamdani Fuzzy Models

(ii) Sugeno Fuzzy Models

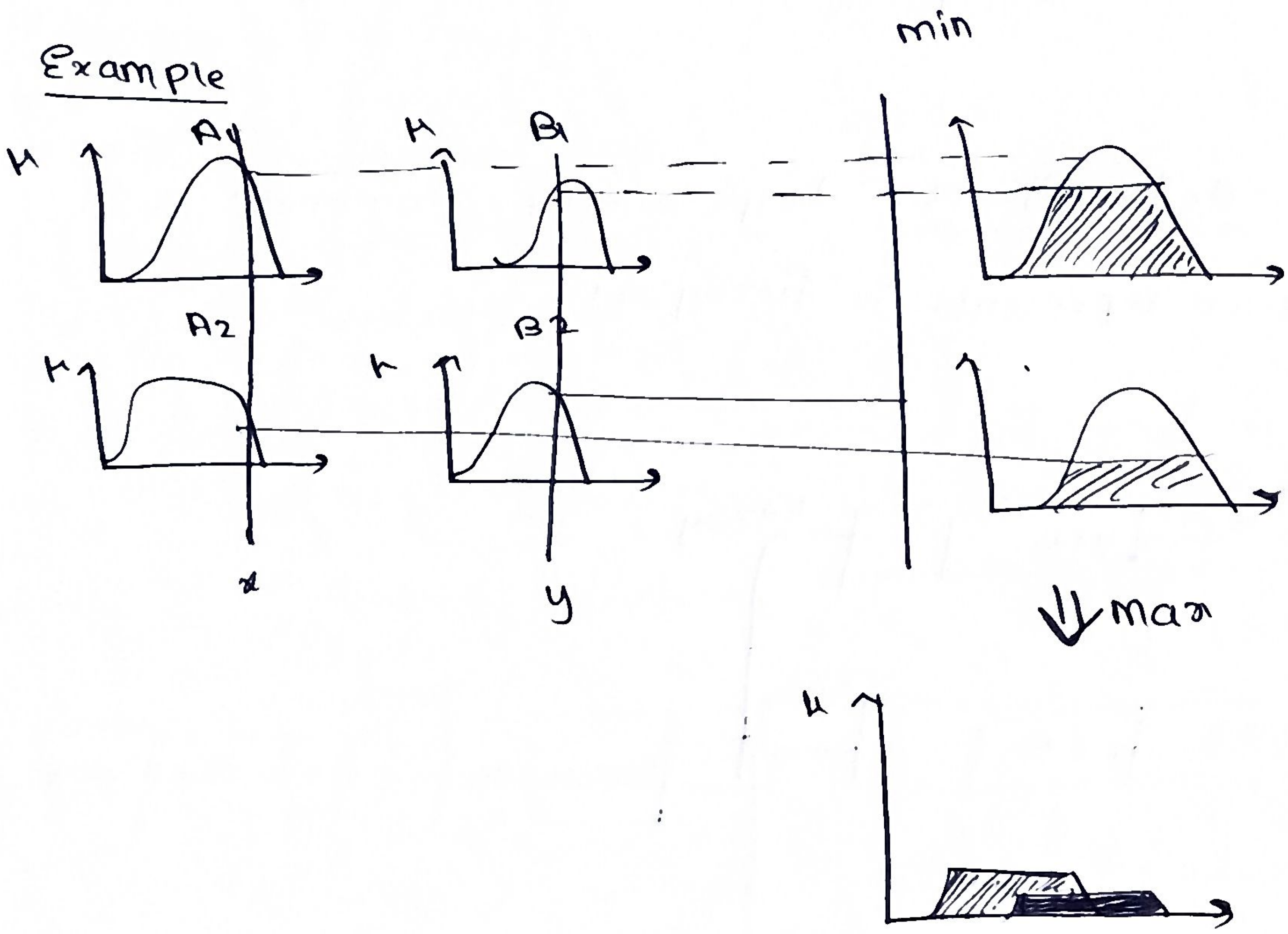
(iii) Tsukamoto Fuzzy Models

* Mamdani Fuzzy Inference Systems

→ introduced as a method to create a control system by synthesizing a set of linguistic control rules, obtained from experienced human operators.

→ General steps - fuzzify inputs, apply a fuzzy operation, apply an implication method, apply an aggregation method & defuzzify

Example



* Application of Mamdani FIS

→ developed as an attempt to control a steam engine & boiler combination

→ Two fuzzy inference systems were used as 2 controllers to generate the heat input to the boiler and throttle opening of the engine cylinder to regulate the steam pressure in the boiler and the speed of the engine.

→ Since the plant takes only crisp values as inputs, we have to use a defuzzifier to convert a fuzzy set to a crisp value.

* Defuzzification

→ Refers to the way a crisp value is extracted from a fuzzy set as a representative value.

→ Different methods include:

(i) centroid of area
(ii) Bisector of area
(iii) Mean of maximum
(iv) Smallest of maximum
(v) Largest of maximum

(i) Centroid of Area

$$z_{COA} = \frac{\int_z \mu_A(z)\, z\, dz}{\int_z \mu_A(z)\, dz}$$

$\mu_A(z)$ = aggregated output MF

→ most widely used strategy

→ similar to calculation of expected values in a probability distribution.

(ii) Bisector of Area

$$z_{BOA} : \int_{\alpha}^{z_{BOA}} \mu_A(z)\, dz = \int_{z_{BOA}}^{\beta} \mu_A(z)\, dz$$

where $\alpha = \min \{ z \mid z \in Z \}$

$\beta = \max \{ z \mid z \in Z \}$

The vertical line $z = z_{BOA}$ partitions the region between $z = \alpha$, $z = \beta$, $y = 0$ and $y = \mu_A(z)$ into 2 regions with the same area.

(iii) Mean of Maximum : average of the maximizing z at which the MF reaches the maximum μ

$$z_{MOM} = \frac{\int_{z'} z\, dz}{\int_{z'} dz'}$$

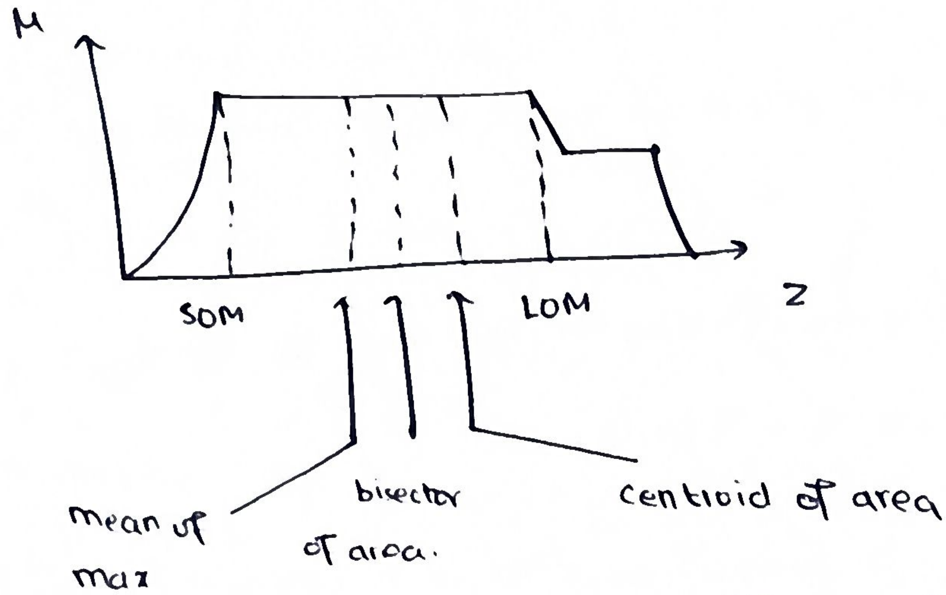
NOTE: The MOM is the defuzzification strategy employed in Mamdani's fuzzy logic controllers.

(iv) Smallest of Maximum (SOM) : z_{SOM} is the minimum in terms of magnitude of the maximizing z.

(v) Largest of Maximum (LOM): z_{LOM} is the maximum in terms of magnitude of the maximizing z.

NOTE: SOM & LOM have biases, not very commonly used defuzzification strategies.

Various Defuzzification Schemes



Look at examples for single input, single output and two input, single o/p Mamdani fuzzy model in TB

* Operations on Mamdani FIS

(i) AND - usually T-norm - calculate firing strengths w/ AND'ed antecedents

(ii) OR - usually T co-norm - calculate firing strengths w/ OR'ed antecedents

(iii) Implication operator - usually T-norm - calculate qualified consequent MFs based on firing strengths

(iv) Aggregate operator - usually T-conorm - agg. qualified consequent MFs to generate overall output MFs

(v) Defuzzification operator - transform an output MF to a crisp single output value.

* Sugeno Fuzzy Models

→ also called the TSK fuzzy model - proposed by Takagi, Sugeno and Kang

→ A typical fuzzy rule in a Sugeno fuzzy model has the form

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

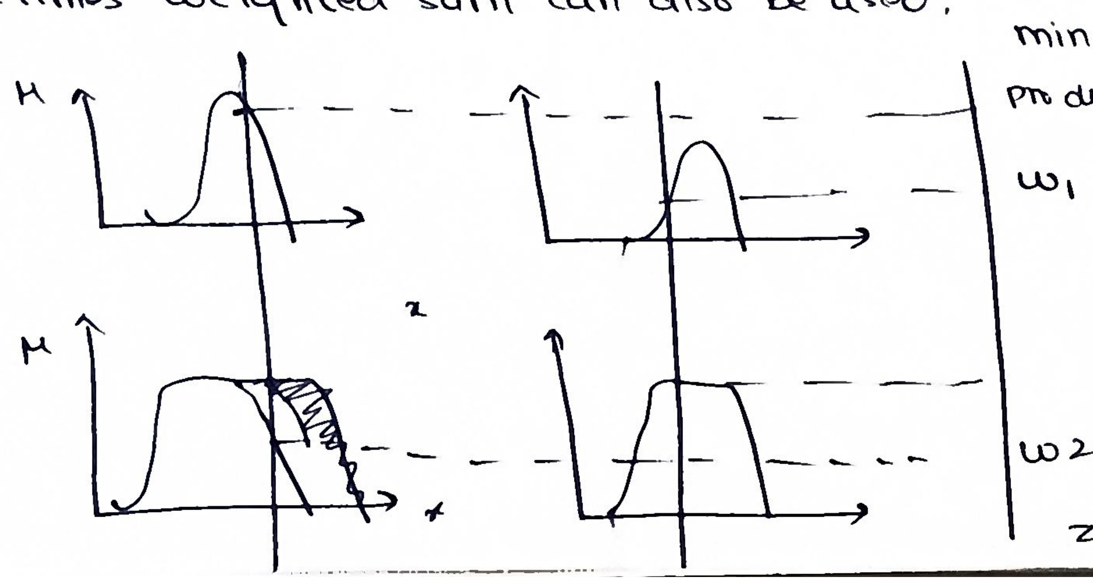
where A and B are fuzzy sets in the antecedent

$z = f(x,y)$ is a crisp function. Usually, $f(x,y)$ is a polynomial in the input variables x and y.

→ When $f(x,y)$ is a first order polynomial, the resulting fuzzy inference system is called a first order Sugeno fuzzy model.

→ When f is a constant, then it is a zero order Sugeno fuzzy models - o/p is a smooth fn. as long as the neighboring MFs in the antecedent have enough overlap.

→ Since each rule has a crisp output, the overall output is obtained via weighted average, thus avoiding the time consuming process of defuzzification required in a Mamdani model. Sometimes weighted sum can also be used.



$z_1 = p_1x + q_1y + r_1$

$z_2 = p_2x + q_2y + r_2$

weighted average

$$z = \frac{w_1z_1 + w_2z_2}{w_1 + w_2}$$

Read single input & two input single o/p sugeno model from TB

* Tsukamoto Fuzzy Models

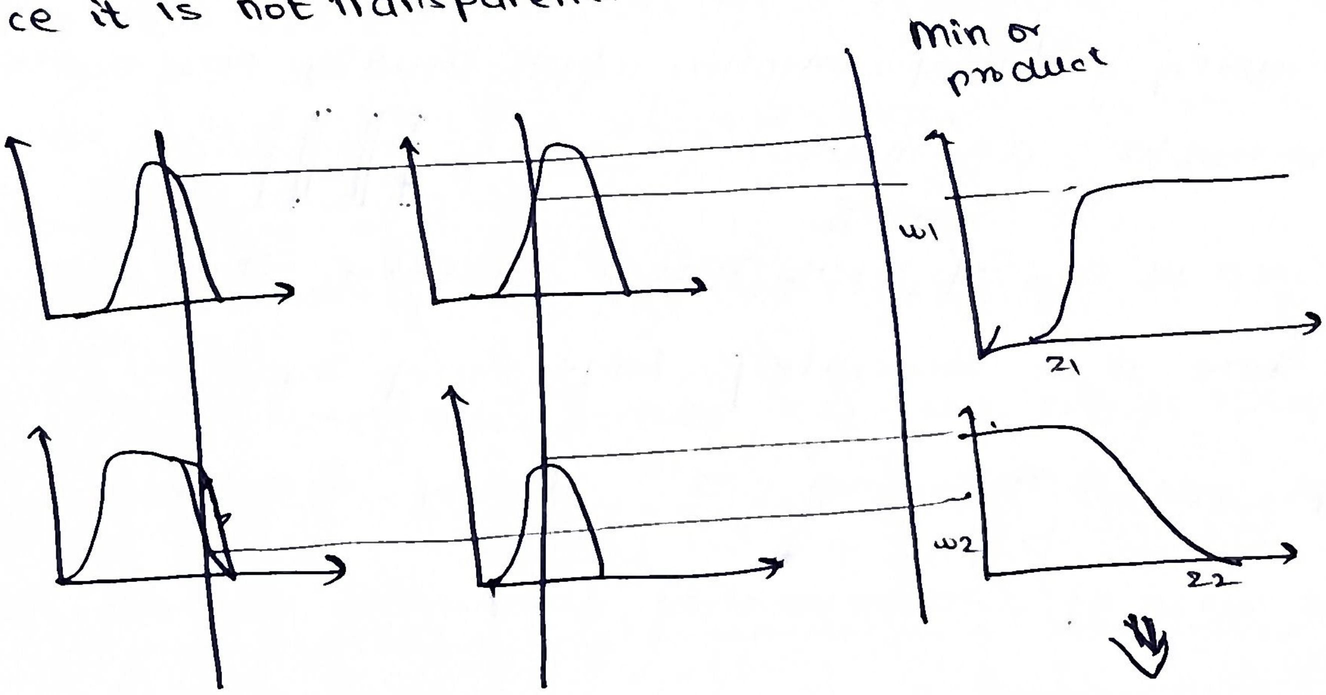
→ Each fuzzy if-then rule is represented by a fuzzy set with a monotonically MF.

→ As a result, the inferred output of each rule is ~~taken as the weighted average of~~ defined as a crisp value induced by the rule's firing strength

→ The overall output is taken as the weighted average of each rule's o/p.

→ It thus avoids the time consuming process of defuzzification

→ Not used as much as Mamdani or Sugeno fuzzy models since it is not transparent.



weighted average

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

Look at example of single i/p Tsukamoto model from TB

* Input Space Partitioning

→ The antecedent of a fuzzy rule defines a local fuzzy region, while the consequent describes the behaviour within the region via constituents.

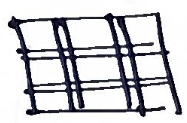
→ The consequent constituent can be a consequent MF from a Mamdani / Tsukamoto fuzzy model, a constant value from a zero order Sugeno model, or a linear equation from a first order Sugeno model.

→ Different consequent constituents result in different fuzzy inference systems, but their antecedents are always the same.

* Methods for Partitioning the Input Space

1 Grid Partition

→ used in designing a fuzzy controller, which usually only involves several state variables as inputs



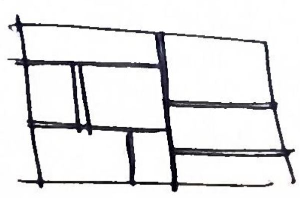
→ needs only a small no. of MFs for inputs ~~MFs~~. However, there are problems when there is a moderately large no. of inputs.

→ eg. w/ 10 i/ps and 2 MFs ⇒ 2^{10} fuzzy if-then rule

called the curse of dimensionality

2 Tree Partition

→ each region uniquely specified along a corresponding decision tree



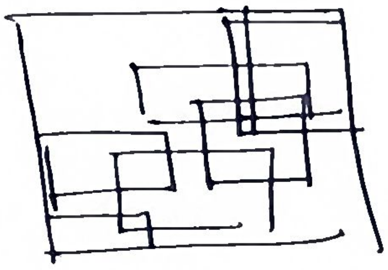
→ relieves the problem of an exponential increase in the no. of rules

→ However, more MFs for each input are needed to define these fuzzy regions, as these MFs usually do not bear clear linguistic meanings like small, big etc.

→ Tree partitioning is used by the CART algorithm

③ Scatter Partitioning

→ A subset of the whole input space is covered, which characterizes a region of possible occurrence of input vectors.



→ limits the no. of rules needed

* Fuzzy Modeling?

→ A fuzzy inference system is usually designed based on the past known behavior of a target system

→ The fuzzy system should then reproduce the behavior of the target system. (eg. a fuzzy expert system for medical diagnosis)

→ Fuzzy modeling has the following features:

(i) incorporate human expertise about target system into the modeling process, ie fuzzy modeling takes advantage of domain knowledge.

(ii) Conventional system identification techniques can be used for fuzzy modeling, ie use numerical data.

* Two Stages in Fuzzy Modeling?

(1) Identification of surface structure

(i) select relevant I/O variables

(ii) choose a type of FIS

(iii) determine linguistic variables associated w/ I/O variables

(iv) design a collection of fuzzy if-then rules

(2) Identification of deep structure (which determines the MFs of each linguistic term)

(i) choose an appropriate family of parameterized MFs

(ii) interview human experts familiar w/ target systems to determine the parameters of MFs used in the rule base.

(iii) Refine the parameters of the MFs using regression & optimization techniques.