$U_i < Q_i$ so $U_i - Q_i$ is -ve

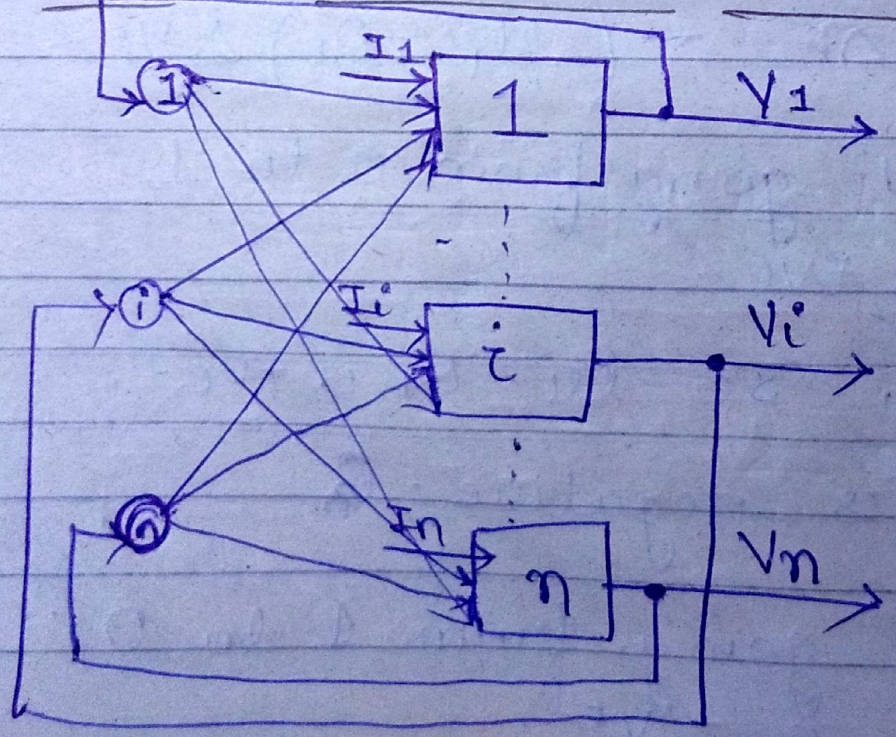
So overall ΔE = -ve.
So still our change in energy is keep on decreasing.

Thus as a result it serves the optimization purpose.

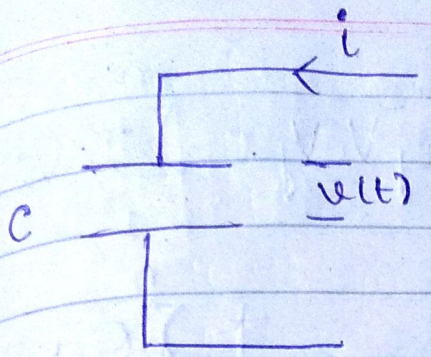
18/9/18.

As the state of the N/W change or remains const. the, energy associated with the N/W decrease → this is the special thing about the Hopfield N/W.

Case II. Continuous Neuron Model.



(Here self-feedback is allowed). i.e, diagonal entry won't be zero.



$$v(t) = \frac{1}{C}\int i\,dt$$

$$\frac{dv(t)}{dt} = \frac{i}{C}$$

u_i is continuously increasing with time so here we will write,

$$\frac{du_i}{dt} = \sum_{j=1}^{n} T_{ij} V_j + I_i \qquad i = 1 \text{ to } n$$

$\left\{ \frac{du_i}{dt} \rightarrow \right.$ what is coming to the node [rate of change of input] $\Big\}$

$u_i \rightarrow$ total input received by a node.

$$\boxed{V_i = g_i(u_i) = \frac{1}{1+e^{-\frac{u_i}{u_0}}}}$$

$\frac{1}{u_0}$ = same as λ (which give gain).

generally have value = 1

(It controls the slope of the sigmoidal function when u_i is close to 0).

$$\Rightarrow E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} V_i V_j - \sum_{i=1}^{n} I_i V_i + \sum_{i=1}^{n} Q_i V_i$$

(But here we are taking threshold = 0 so $\sum Q_i V_i) = 0$

Now, $E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} V_i V_j - \sum_{i=1}^{n} I_i V_i$

$$\frac{dE}{dt} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} \left\{ \frac{d(V_i \cdot V_j)}{dt} \right\} \quad T_{ij} \approx w_{ij} \}$$

$$- \sum_{i=1}^{n} I_i \frac{dV_i}{dt}$$

$$\Rightarrow \frac{dE}{dt} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} \left(\frac{dV_i}{dt} \cdot V_j + V_i \frac{dV_j}{dt} \right)$$

$$- \sum_{i=1}^{n} I_i \frac{dV_i}{dt}$$

$$= -\sum_{i=1}^{n} \left[g_i'(U_i) \cdot \frac{dU_i}{dt} \right] \cdot \left[\frac{dU_i}{dt} \right]$$

$$\boxed{\Rightarrow \frac{dE}{dt} = -\sum_{i=1}^{n} g'(U_i) \cdot \left[\frac{dU_i}{dt} \right]^2}$$

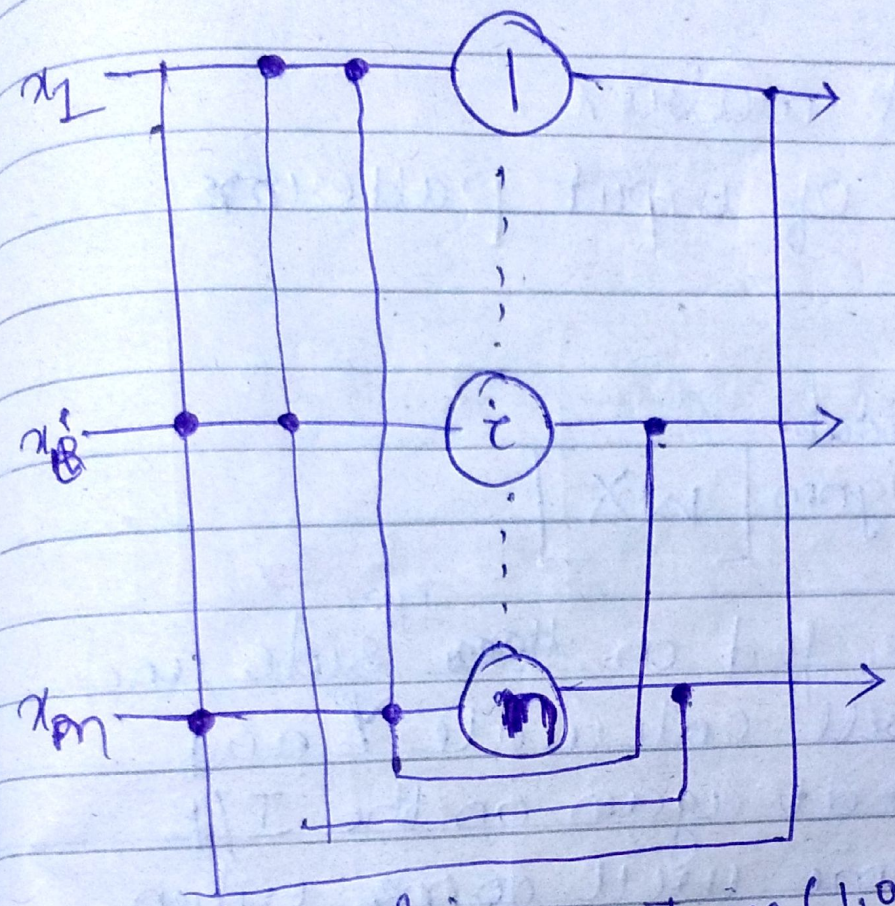
When $\frac{dE}{dt}$ remains const., that is the optimum ~~term~~ state for the Hopfield network & that is our optimum hopfield network (or optimum design of the product)

$\left[\frac{dU_i}{dt} \right]^2$ = always remains positive,

$g'(U_i)$ = increasing function (always goes on increasing from -ve to +ve).

So $\frac{dE}{dt}$ = will always be -ve and decreasing.

Auto - associative Hopfield network



→ have binary entries (1, 0 or -1)

$Y[1], Y[2], \ldots Y[m] \ldots Y[M]$ } required to be stored

→ It is meant to store certain fundamental patterns (fundamental memory).

When any of these pattern (in corrupted form) is presented in this input side, it should able to retrieve the exact pattern or fundamental memory that is stored in the network.

$$= \sum_{m=1} [Y_m][Y_m]^T - M[I]$$

I, unit matrix.
M = no:- of input patterns

$$= \text{sigm}[WX]$$

this Y is fed on ~~the~~ I/P side and
n we will calculate Y and
ill feed that again on the I/P
e, a time will come when
ee will be no refinement
d that will be our final
ndamental memory, and it
hould be very similar to one
f the stored fundamental
memory.

$$Y[1] = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} ; \quad Y[2] = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

$$x = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$W = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Calculated $Y = \text{sign}\left(\begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \right)$

$$= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$$

Calculated $Y = \text{signm}\left(\begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right)$

$= \text{signm}\left(\begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}\right)$

In hopfield only,

signm(0) = means O/P will remain unchanged (same as I/P).

signm(>0) = 1 (output = 1)

signm(<0) = −1 (Output = −1)

generally, sign(0) = 0.

sign(>0) = 1 sign(<0) = −1

In hopfield sign(0 and >0) = 1, sign(<0) = −1

$X = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ $\qquad Y = \text{signm} \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$

This Y is feedback to the input, then answer is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

signm(0) = same as X = 1

" = same as X = 1 $\qquad$ so O/P = $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

sign(4) = 1

This is same as Y[1]

$\boxed{\text{So our } X \text{ pattern is} \downarrow Y[1] \text{ same as} \uparrow}$

20/9/18

Radial Basis Network :-

→ It is also used for function approximation.

→ Unsupervised learning as well as supervised learning

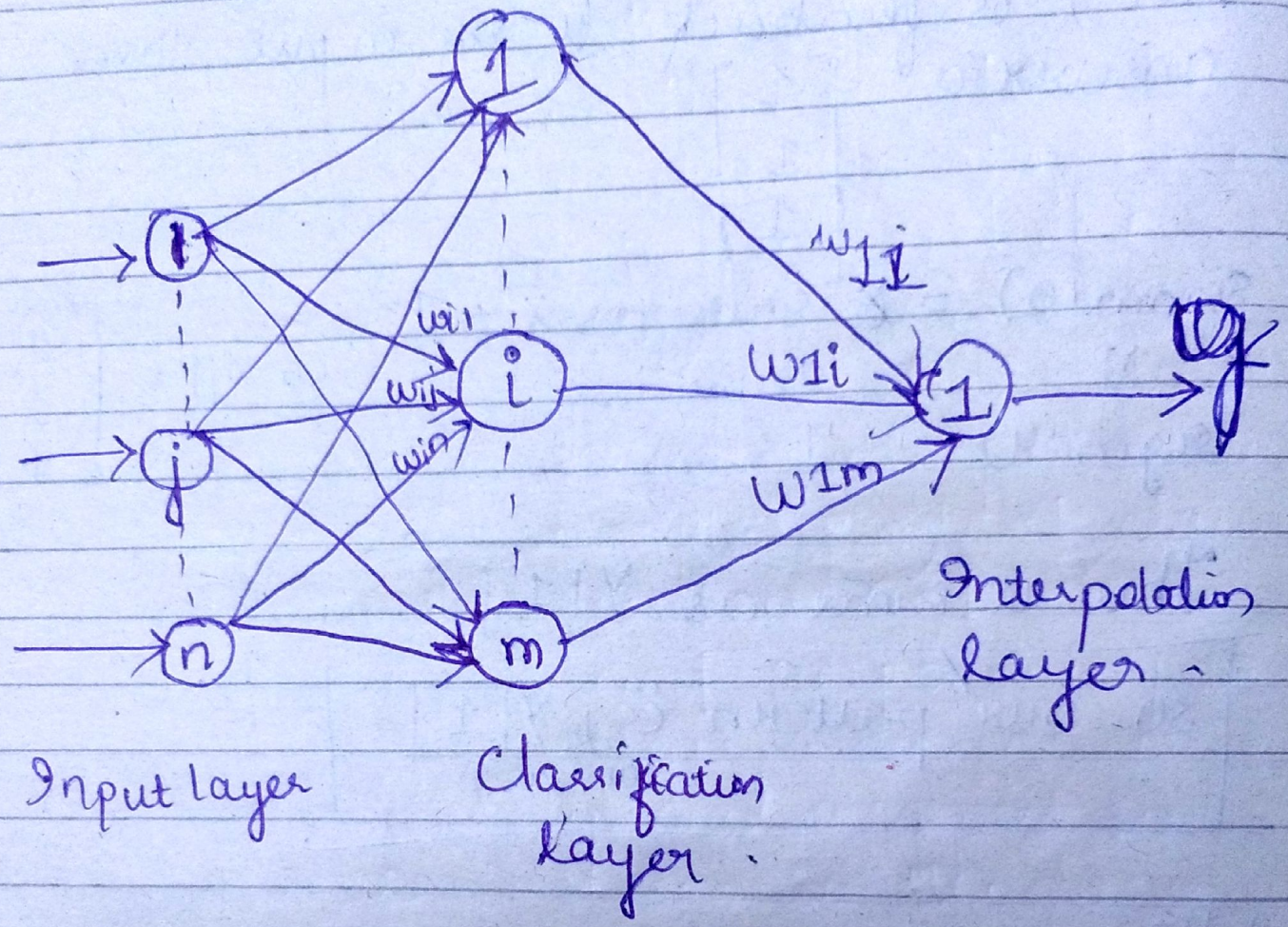
→ It is a hybrid kind of network.

→ Feed forward network.

→ There are two types of layer :-

a) Classification layer

b) Interpolation layer.



$$y = f(x_1, x_2 \cdots x_n).$$

→ If you pick up given presentation or input pattern, then we will see in which class the input layer will fall and accordingly our output will be calculated.

→ Set of I/P patterns is need to be classified to one of the class in classification layer

so this there are 2 stages :-

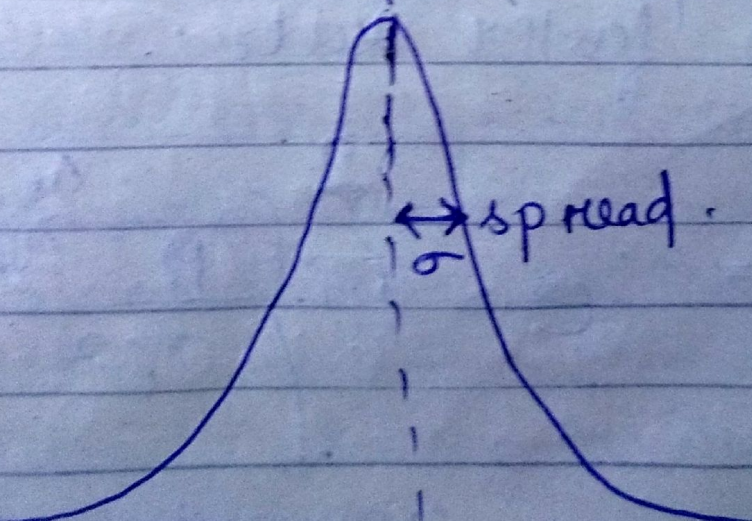
At stage 1 → classification of input patter

At stage 2 → the interpolation of y wrt to input variables

→ If some input pattern is a member of i is highest then to a larger extent the class i will affect the value of y.

If some input pattern is not all a member of any class, then that class won't affect the value of y at all.

→ Because the activation functions used in the classification units are gaussian / radial basis function. That is why it is known as radial basis network.



~~If the distance of the class~~
weight vector is representing the centre of the particular class in classification layer.

If the given presentation is close to the centre of the particular class, then activation is maximum. (and the response is maximum).

If the ~~spread is maximum~~ given presentation is not close/far of the partical class, then activation is less and hence spread is more and response is poor.

Now when the spread is ~~more~~ less, then the output gradually decrease but if spread is ~~too~~ more, then the output vary (deviation is more) at a faster rate.

$$O_i = e^{-\sum_{j=1}^{n} \frac{(x_{pj} - x_{ij})^2}{\sigma_i^2}}$$

centre

components of input presentation

Components of class i or wt vector related to class i

→ It has also got 2 phases:
① Training phase.
② Testing phase.

→ Input data used for training

set of p i/p pattern

$X[1] \longrightarrow Y[1]$ (scalar)

$X[2] \longrightarrow Y[2]$ (scalar)

$\vdots$

$X[p] \longrightarrow Y[p]$

$\vdots$

$X[P] \longrightarrow Y[P]$

→ Training phase
① Classification phase.
② Interpolation phase.

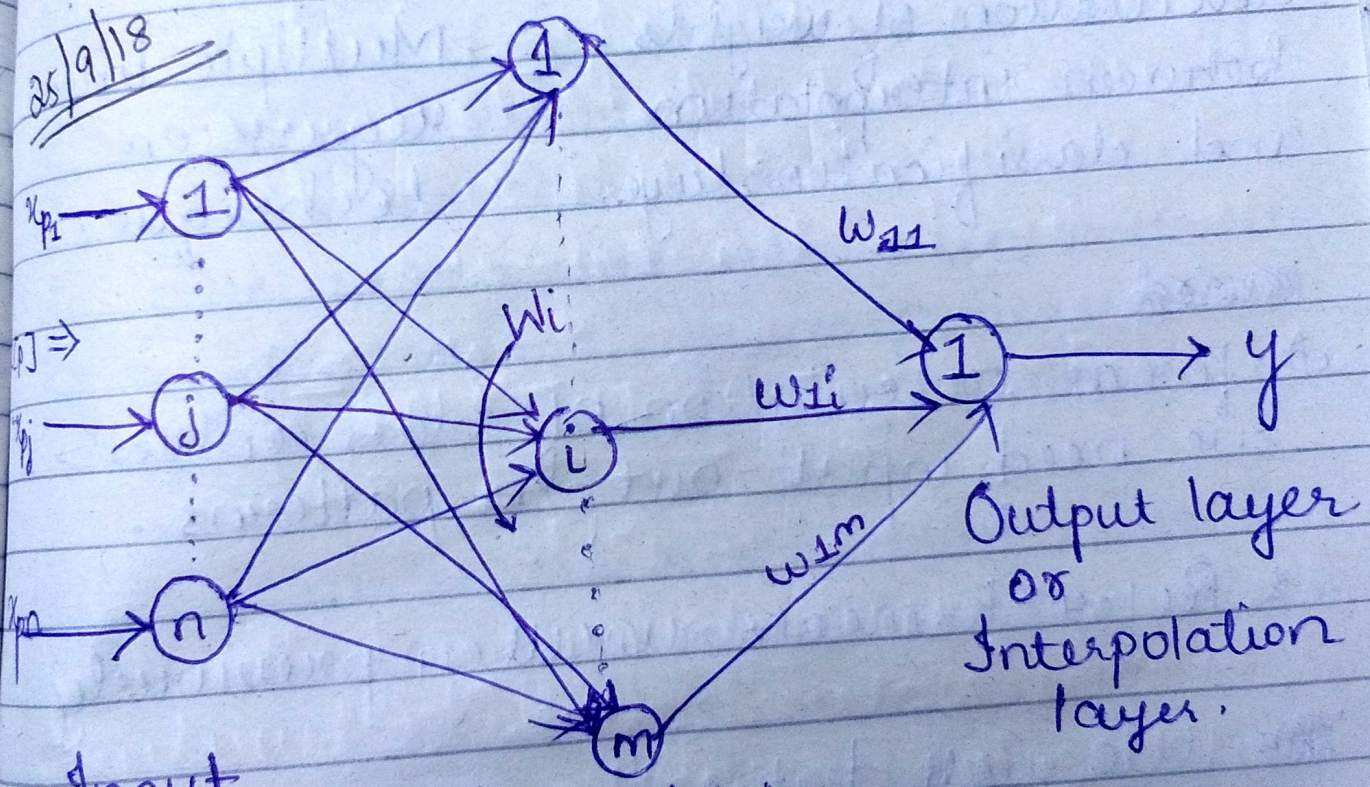
→ In classification, our objective is to find the centre of all the classes.

→ we are looking out the final values of connection weights b/w classification & Interpolation layer

① Classification → k-means method.
② Class centres are to be finalized using nearest neighbourhood method.

→ The ~~pattern~~ eucledian distance b/w the input pattern and all the class centres are calculated.
the one with the minimum eucledian distance is declared as winner node and that input pattern will be mapped to that class or classification node.

25/9/18



Radial Basis fn. Network

Class centres → K-means method.
(or any of the clustering algorithms can be used).

Calculation of spread or width of RBF, σ – Nearest neighbourhood method.

Calculation of weights between interpolation and classification layer – Multiple linear regression technique.

As far as calculation of class centres → we need input-output patterns.

* Refer K-means written previously.

We will take the mean of the all the input patterns that is mapped to one-node and this mean value gives the final wt for that node for that iteration.

③ → Calculation of weights → multiple regression technique.

K-means

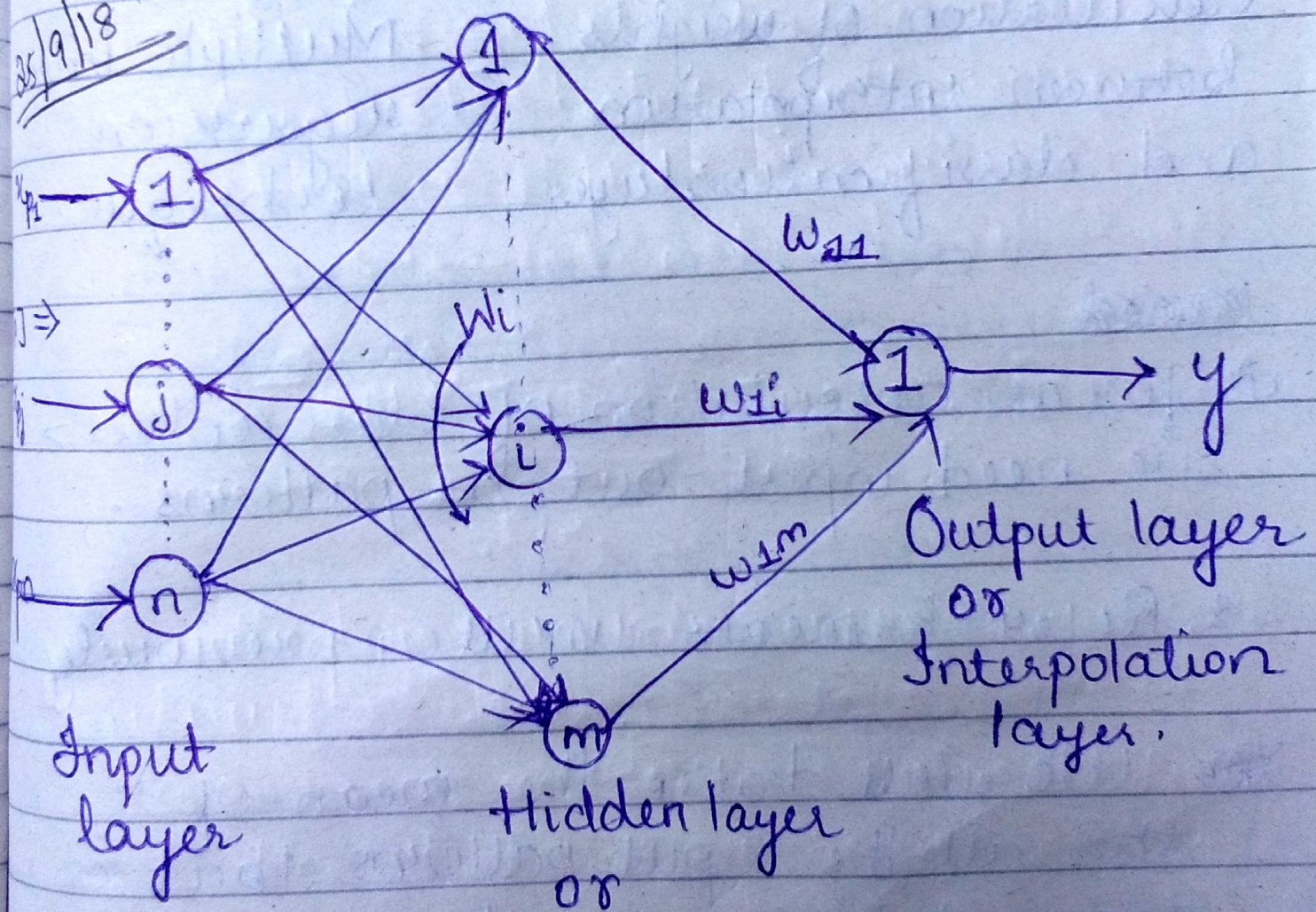
→ Randomly we will pick up input patterns and will initialize the node in the classification layer. (so we have to pick the same no:- of nodes in classification layer as input patterns).

→ We will find that how many patterns have been mapped to one class and will take average of their values as the centre of the class (update the class centres).

→ When we are getting same mapping pattern again and again, then our training for classification is finished

→ The ~~pattern~~ eucledian distance b/w the input pattern and all the class centres are calculated.
the one with the minimum eucledi-an distance is declared as winner node and that input pattern will be mapped to that class or classification node.

25/9/18



Radial Basis fn. Network

Class centres → K-means method.
(or any of the clustering algorithms can be used).

Calculation of spread or width of RBF, σ – Nearest neighbourhood method.

Calculation of weights between interpolation and classification layer – Multiple linear regression technique.

As far as calculation of class centres → we need input-output patterns.

* Refer K-means written previously.

We will take the mean of the all the input patterns that is mapped to one-node and this mean value gives the final out for that node for that iteration.

Now calculation of spread of RBF fn. for the ith unit in the Hidden layer.

$$\left[\sum_{j=1}^{n} (\hat{x}_{kj} - \hat{x}_{ij})^2\right]^{1/2} = \sigma_i$$

$$\sigma_i = \left[\frac{1}{p}\sum_{k=1}^{p}\sum_{j=1}^{n} (\hat{x}_{kj} - \hat{x}_{ij})^2\right]^{1/2}.$$

We calculated the ecludein distance ~~all the~~ of ith unit from p nearest neighbours. and then averaging.

$\hat{x}_{ij}$ = class centre associated with i comprising of n components.

* We have just done the averaging of distance of the p-nearest neighbours of ith node.

And taking the square root it gives σ_i (spread for ith node).

p = design parameter.

W_i = final class centres or wt. vectors associated with i which is calculated by k-means.

~~p = design~~

→ p will always be less than m.

<u>Calculation of weights between Hidden layer and output layer</u>

→ Once we got the value of the class centres and spread, we will present the first input pattern and calculate the activation function for all the nodes in the hidden layer.

$$a_i = e^{\left(\frac{-\sum_{j=1}^{n}(x_{pj} - \hat{x}_{ij})^2}{\sigma_i^2}\right)}$$

The one ~~with~~ which has got the highest value of the activation, that node's wt vector is nearest to the given presentation.
(The one which has got the highest value of activation will be declared as winner node).

→ Present 1st pattern and calculate activation for all nodes, then second and again calculate the activation and so on upto Pth presentation.

$$\underset{W = L \times m}{\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1i} & \cdots & w_{1m} \\ \vdots & & & & & \\ w_{l1} & w_{l2} & \cdots & w_{li} & \cdots & w_{lm} \\ \vdots & & & & & \\ w_{L1} & w_{L2} & \cdots & w_{Li} & \cdots & w_{Lm} \end{bmatrix}} \underset{A = m \times P}{\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{p1} & \cdots & a_{P1} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{1i} & a_{2i} & & a_{pi} & \cdots & a_{Pi} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{1m} & a_{2m} & & a_{pm} & \cdots & a_{Pm} \end{bmatrix}}$$

Output will be of dimension $L \times P$.

a_{11} = activation of 1st node for 1st presentation.

a_{1i} = activation of ith node for " "

$w_{11}, w_{12} \cdots w_{1m}$ are the weights of 1st output node for n nodes in Hidden layer.

$w_{l1}, w_{l2} \cdots w_{lm}$ → are the weights of the lth output node.

* (Here in calculation of W we are considering that we have L nodes in o/p layer).

But most of time we will come across 1 o/p.

$$y_{1p} = \sum_{i=1}^{m} w_{1i} \cdot a_{1i} \qquad p = 1 \text{ to } P$$

$$\vdots$$

$$y_{Lp} = \sum_{i=1}^{m} w_{li} \cdot a_{li} \qquad l = 1 \text{ to } L$$

Distance between targets and [W][A] should approach to 0.

W is not known to us.
We have to calculate W.

$$\| [T] - [W][A] \| \Rightarrow 0$$

↳ Target outputs.

We know T and A.

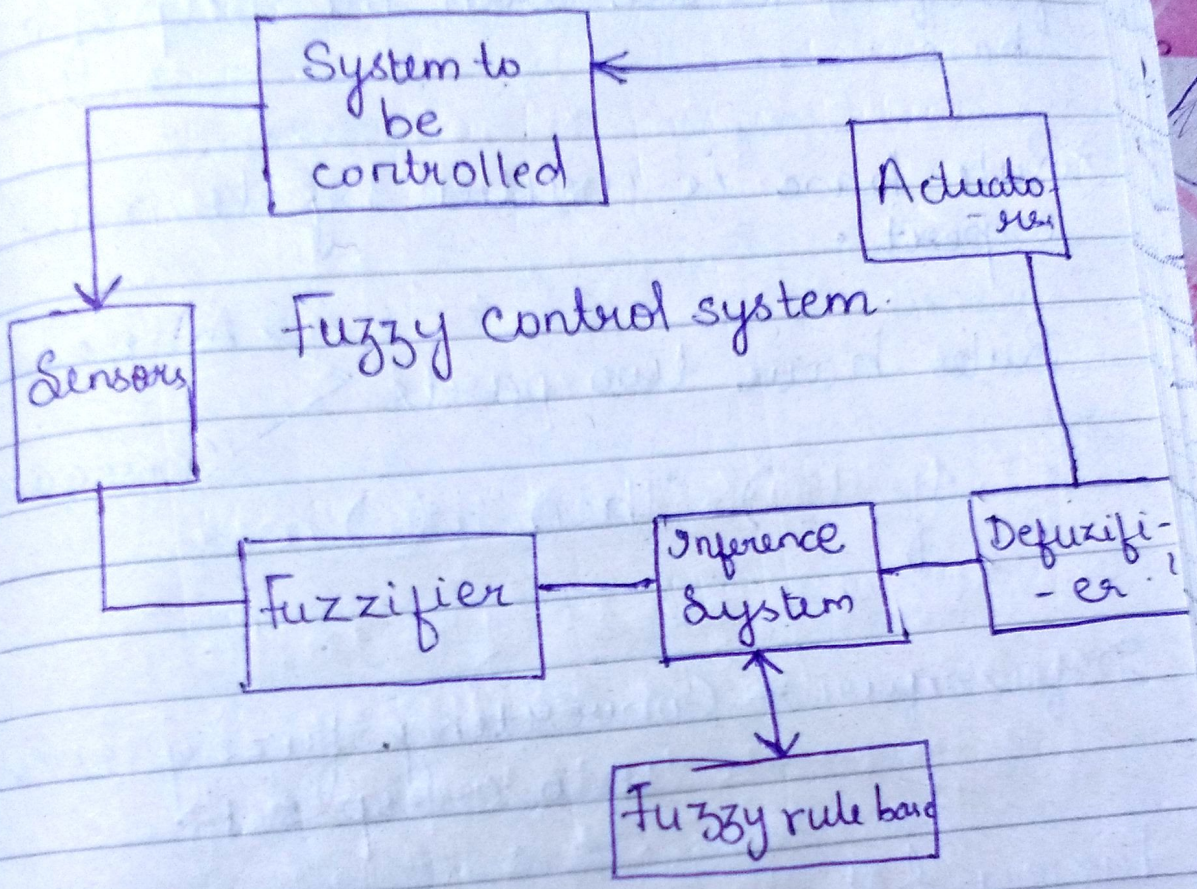
$$\boxed{[W] = [T][A]^{-1}}$$

RBF^n steps ←

Calculate class centre
↓
get value of spread
↓
Calculation of activation
↓
Calculation of wt matrix

21/9/18

Fuzzy Systems



→ In case of fuzzy control system → algorithms model.

In case of ~~mathematical~~ conventional system → some mathematical model.

Initially we have crisp info which is fuzzified with the help of fuzzifier

Sensors are used to detect the deviation from the critical system.

Inference System:

Draw some inference from the fuzzified value with the help of rule base.

→ Rule base is written by domain expert.

Rules have two parts < Antecedant, Consequent.

If ( ) Then ( ).

Defuzzifier – Converting fuzzy sets into crisp set.

Fuzzy Set Theory:-

$X = [a, b, c, d, e, f, g]$.

↳ Universal set.

$A = [a, b, c]$ (Crisp set – either member or not)

If you pick any ~~mend~~ element from X then it is either member of A or not.

→ A is being defined over the universal set.

Let B be another set defined over the universal set.

$B = [c, d, e]$ (this is also a Crisp set).

<u>Set operations which can be performed over the crisp sets</u>.

① Union ② Intersection
③ Complement ④ Difference.

Now the idea of crisp set is extended to the fuzzy set.

In crisp set → degree of membership is either 1 or 0.

In fuzzy set → degree of membership can be b/w 0 & 1 (including).

→ represented as $\tilde{F}$ or $\underset{\sim}{F}$ or F.

$\tilde{F} = \{0.7/a, 0.5/b, 0.4/c, 0.9/d, 1/e, 0/f, 0.5/g\}$.

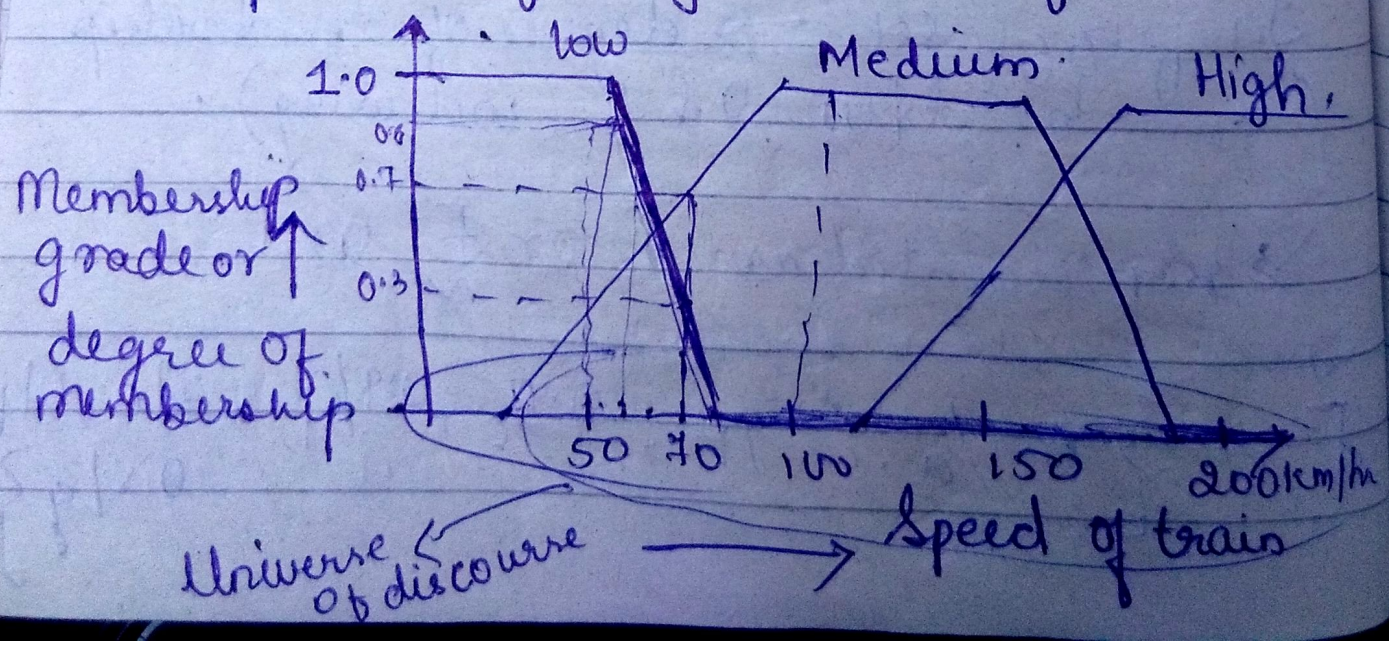
The membership grade of a is 0.7
" " " " b is 0.5
" " " " c is 0.4
" " " " d is 0.9
" " " " e is 1
" " " " f is 0
" " " " g is 0.5

$\mu_F(x) = 0.6$ → This means membership grade of x in fuzzy set F is 0.6 (or x to fuzzy set F).

x → one of element picked from Universal set.

$$\mu_F(x) = \frac{1}{1+x^2} \quad \text{(Membership function)}.$$

We represent fuzzy sets as follow:-



* Medium, Low, High are all fuzzy sets defined over the universe of discourse.

→ The whole universe of discourse is divided into some liguistic labels or fuzzy set (defined over the universe of discourse).

$\mu_{LOW}(x)$ = This is to be decided by the fuzzy set low.

The speed of train is covered by 3 fuzzy sets → low, medium and high.

These fuzzy sets taken together forms the term set.

If take 100, then its membership grade to medium = 1 and low is 0 and high is 0.

For 70, then its membership grade to medium is 0.7, low = 0.3 and high = 0.

→ It all depends upon the perception of the designer → what his perception about the fuzzy sets. (This gives the design of the fuzzy sets).

→ Here we deal everything wrt domain expert.

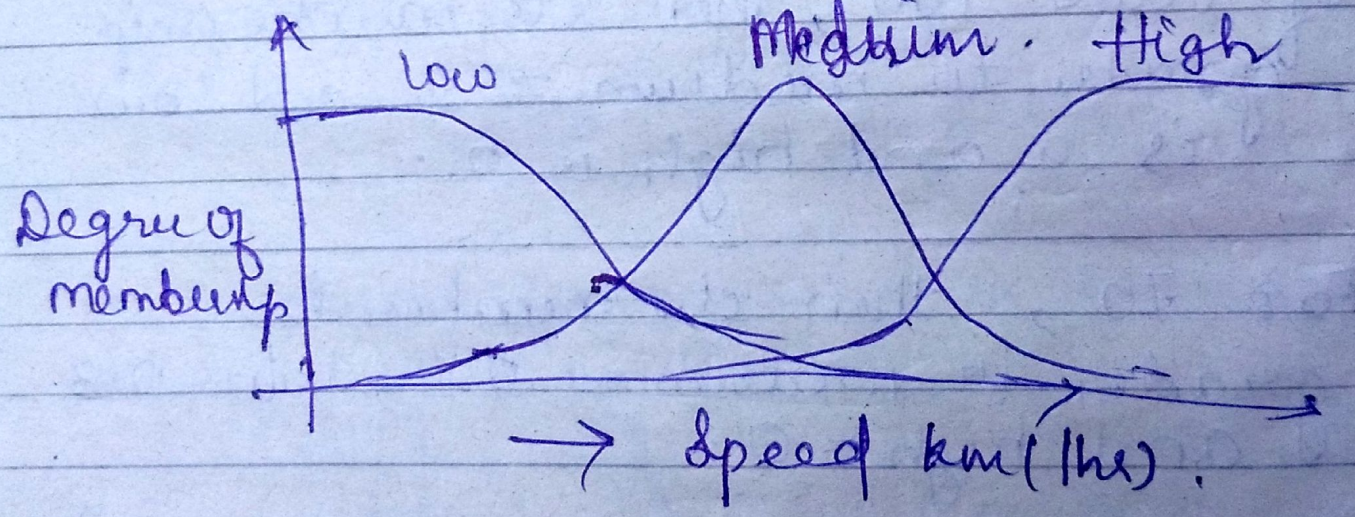
→ Domain experts always talk in terms of fuzzy domain so that is why we move from crisp domain to fuzzy domain to use the rule base which is in fuzzy domain.

→ $\mu_F(x) = 0.8$

$\Downarrow$

50

→ Medium could also be in the form of gaussian function as shown below and also high & low.



→ The no. of fuzzy sets → depends upon the tuning of the fuzzy system.

→ The no:- of fuzzy sets → gives better performance when fuzzy system is more close to the classical system.

$X = \{a, b, c, d, e, f\}$.

$A = \{0.5/a, 0.3/b, 0.9/c, 0.4/d, 0.2/e, 0.5/f\}$.

→ fuzzy set A defined over X.

$B = \{0.3/a, 0.9/b, 0.7/c, 0.3/d, 0.9/e, 0.6/f\}$.

If we want to calculate $A \cup B$.

= max of $\mu_A(x)$ and $\mu_B(x)$

$$\boxed{A \cup B = \max\{\mu_A(x), \mu_B(x)\}}.$$

$A \cup B = \{0.5/a, 0.9/b, 0.9/c, 0.4/d, 0.9/e, 0.6/f\}$.

$$\boxed{A \cap B = \min\{\mu_A(x), \mu_B(x)\}}.$$

$\boxed{\bar{A} = \{ 1 - \mu_A(x) \}}$ for all $x \in X$.

$\bar{A} = \{ 0.5/a, 0.7/b, 0.1/c, 0.6/d, 0.8/e, 0.5/d \}$.

$\boxed{\bar{B} = \{ 1 - \mu_B(x) \} \ \forall \ x \in X}$

Core of a fuzzy set.

→ Core of a set is a crisp set comprising all those elements whose membership grade is 1.

Core of A is an empty set.
" " B is an " "

} because no elements are there with memb. grade is 1.

α-cut of a fuzzy set.

→ α can take any value from 0 to 1.

α-set should always be a crisp set comprising of all elements whose membership grade is $\geq \alpha$.

If $\geq \alpha \rightarrow$ then soft α-cut.

If $> \alpha \rightarrow$ then hard α-cut.

28/9/18

TUTORIAL - 4.

Q. Design a BAM network to associate following letters with simple bipolar codes.

```
• # •        • # #
# • #        # • •
# # #        # • •
# • #        # • •
# • #        • # #
```

$\updownarrow$ $[-1, 1]^T$ $\qquad$ $\updownarrow$ $[1, 1]^T$

= 1
• = -1

These are the two pairs that are to be stored as fundamental memory in BAM (Bipolar associative memory).

Input vector contains 15 enteries.
Output vector will be either [-1, 1] or [1, 1]

(1) Find the recall of the N/W if it is fed by one of the input/output patterns.

(2) Calculate response of the network if

it is fed by $[0\ 1]^T$ on the output side.

3) Find the equilibrium point ~~of~~ ~~its~~ it is fed by $[0\ 1]$ on o/p side and

$$W = \sum_{p=1}^{P} X[p]\, Y[p]^T$$

$$[0\ 0\ -1\ 1\ 1\

$$W = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ -1 & -1 \\ -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$W = \begin{bmatrix} 0 & -2 \\ 0 & 2 \\ 2 & 0 \\ 0 & 2 \\ 0 & -2 \\ -2 & 0 \\ 0 & 2 \\ -2 & 0 \\ -2 & 0 \\ 0 & 2 \\ 0 & -2 \\ -2 & 0 \\ -2 & 0 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}$$

① Feeding 1st pattern in the Input side

$$\begin{bmatrix} \\ W^T \cdot X \end{bmatrix}^T = \begin{bmatrix} -14 & 16 \end{bmatrix}$$

$$Y' = sig[X^T W]$$

$$= \begin{bmatrix} -1 & 1 \end{bmatrix}$$

$$Y^T = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$ → Same as the given O/P pattern.

Feeding above Y pattern O/P.

$$X = sig\begin{bmatrix} \\ W Y \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$ → I/P Pattern 1. (✓)

$$Y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

② $\left(\text{sign}\left(W \times \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \right) = X$.

$$= \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

{ Here sign(0) = 0 becaus in I/P we are assuming all the entries are 0 }

③ ~~Now as we know~~

By providing some additional informatio we are able to arrive at some fundamental memory.

Now here some corrupted I/P pattern is also given so wherever in ② we have got 0 we can see I/P pattern and write it as it is (if there is a disagreement between the 2 I/P pattern else it will be same as I/P pattern i.e., 2 pattern)

[0 0 -1 0 0 1 0 1 1 0 0 1 1 -1 0

$X^T = [-1\ 1\ 0\ 1\ -1\ 0\ 1\ 0\ 0\ 1\ -1\ 0\ 0\ 0\ 1]$

Calculated in Q(2).

Now we have something on the I/P side so where sign(0) = 0 we will write the I/P given in ques as it is.

So now our calculated X is:-

[-1 1 -1 1 -1 1 1 1 1 1 -1 1 1 -1 1]

Now due to the additional I/P pattern we are able to retrive the exact fundamental memory.

Here it matches with the 1st I/P pattern or fundamental memory.

01/10/18

Some more terms related to fuzzy set theory :-

Let A be a fuzzy set defined over universe of discourse X.

$$X = \{a, b, c, d, e, f\}$$

$$A = \{0.2/a,\ 0.7/b,\ 0.5/c,\ 0.9/d,\ 0.2/e,\ 0.5/f\}$$

(OR)

$$A = \{0.2/a + 0.7/b + 0.5/c + 0.9/d + 0.2/e + 0.5/f\}$$

Core (A) → comprising of all those elements whose membership grade = 1. It is also a crisp set.

Support (A) = also a crisp set. All those elements whose membership grade is greater than 0.

$$\text{Support}(A) = \{a, b, c, d, e, f\}.$$

Height of a fuzzy set :- The maximum value of the membership grade.

$$\text{Height}(A) = 0.9.$$

$\bar{A}$ or negation of $A = \{ 0.8/a, 0.3/b, 0.5/c, 0.1/a, 0.8/e, 0.4/f \}$

$= \{ 1 - \mu_A(x) \ \forall x \in X \}$.

$\bar{\bar{A}} = A$.

<u>Dilation of a fuzzy set :-</u>

$Dil(A) = \{ [\mu_A(x)]^{1/2} \ \forall x \in X \}$

<u>Concentration of a fuzzy set :-</u>

$Con(A) = \{ [\mu_A(x)]^2 \ \forall x \in X \}$

<u>Hot</u> :- It is a fuzzy term / linguistic level / linguistic value. ~~is used to measure~~

Suppose 50°C → what is the degree of extent of 50° to be called as hot.

Eg :- I can say 50°C is hot with a membership grade of 0.6 as per

the perception of a particular designer

Now 50°C → what is the degree of extent of 50° to be called as a very hot.

So the same 50°C to be called very hot will be less than 0.6,
$= (0.6)^2 = 0.36$.

Now, very very hot $= (0.6)^4$

This makes the membership grade less and less by increasing the concentration or just finding the concentration of fuzzy set.

<u>Term set</u> :- Set of fuzzy sets used to represent a particular universe of discourse.

To represent temperature we can have fuzzy sets low, Medium, High and Very High over some universe of discourse (for eg 0-100°C).

So all these fuzzy sets / fuzzy level

taken together form the term set.

Suppose there is another fuzzy set $\tilde{B}$ defined over X. (the same universe of discourse) comprising of 6 elements.

$$\tilde{B} = \{0.4/a, 0.6/b, 0.9/c, 0.4/d, 0.8/e, 1/f\}$$

Bonded Sum, $\oplus$

$A \oplus B$ = another fuzzy set.

$$= \{0.6/a, 1/b, 1/c, 1/d, 1/e, 1/f\}$$

Bonded sum of A and B

$$\min(1, (\mu_A(x) + \mu_B(x))) \text{ for } \forall x \in X.$$

Bonded difference $\ominus$

$$A \ominus B = \max\{0, (\mu_A(x) - \mu_B(x))\}$$

$$A \ominus B = \{0/a, 0.1/b, 0/c, 0.5/d, 0/e, 0/f\}$$

Normal fuzzy set

It is the one where atleast one of the ~~fuzzy set~~ membership grade is 1.

So B is normal fuzzy set.

If it is not normal then it is called sub-normal fuzzy set.

So, A is a sub-normal fuzzy set as it has no membership grade = 1.

If one of the membership grade is other than 1 or 0 is called proper fuzzy set.

Power set of a set

$A = \{1, 2, 3\}$.

Power set of A = $\{\{\phi\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}$.

Cartesian product of two sets

$P = \{1, 2, 3\}$. $Q = \{a, b\}$.

$P \times Q = \{(1,a), (1,b), (2,a), (2,b), (3,a), (3,b)\}$.

Cardinality of a set :- The no:- of elements in the set.

Algebraic sum.

$$A + B = \{\mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x)\}$$

$$\forall x \in X.$$

Algebraic Product:

$$A \cdot B = \{\mu_A(x) \cdot \mu_B(x) \ \forall x \in X\}.$$

$$A - B = A \cap \bar{B}$$

$$B - A = B \cap \bar{A}$$

Properties of fuzzy sets.

(1) Commutativity $A \cup B = B \cup A$
$A \cap B = B \cap A$

(2) Associativity $A \cup (B \cup C) = (A \cup B) \cup C$
$A \cap (B \cap C) = (A \cap B) \cap C$

(3) Distributivity:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

04/10/18

Properties of Fuzzy Sets: -

(4) Idempotency

$$A \cup A = A \quad ; \quad A \cap A = A$$

(5) If A is a subset of B i.e, $A \subset B$
then $B = A \cup B$; $A = A \cap B$.

(6) $A \cap \phi = \phi$ (7) $A \cup \phi = A$

(8) $A \cap X = A$ where X is the universe of discourse

(9) $A \cup X = X$.

(10) $A \subset B$ and $B \subset C$ then $A \subset C$

(11) Demorgan Theorem

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

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

(12) $\bar{X} = \phi$

$\bar{\phi} = X$

Q. There are two universe of discourse.

(1) Diseases (Typhoid (T), Viralfever (V), running nose (R).

(2) Symptoms (cough (C), high temp (H), Shivering (S))

Dise / Symp	T	V	R
C			
H			
S			

Their cartesian product, Diseases X Symptom

$= \{(T,C), (T,H), (T,S), (V,C), (V,H), (V,S), (R,C), (R,H), (R,S)\}$.

If two elements are not closely related then that relation may or may not be present.

(1) Classical relation – either they will be connected/related or not ie, 0 and 1

(2) Fuzzy relation:- All the element in the Cartesian product have membership grade b/w 0 and 1.

$R = \{ 0.2/(CT, C), \ldots\ldots, 0.9/(V, H) \}$.
If a pair is strongly related then degree of membership is close to 1. e.g., (V, H).

Let us consider two universe of discourse

$P = \{ P_1, P_2, P_3, P_4 \}$ — Variety of plants

$D = \{ D_1, D_2, D_3, D_4 \}$ — Diseases which can affect plants.

$S = \{ S_1, S_2, S_3, S_4 \}$ → Symptoms.

We are finding fuzzy relation which is connecting P and D :-

ie,

$$R = \begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \end{array} \begin{array}{c} \begin{array}{cccc} D_1 & D_2 & D_3 & D_4 \end{array} \\ \begin{bmatrix} 0.6 & 0.6 & 0.9 & 0.8 \\ 0.1 & 0.2 & 0.9 & 0.8 \\ 0.2 & 0.3 & 0.4 & 0.8 \\ 0.9 & 0.8 & 0.1 & 0.2 \end{bmatrix} \end{array}$$

This contain degree of membership which plant is how much prone to a particular disease.

S = Another fuzzy relation tells what are symptoms of a particular plant disease.

$$
= \begin{array}{c} \\ D_1 \\ D_2 \\ D_3 \\ D_4 \end{array}
\begin{array}{c}
\begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\
\begin{bmatrix}
0.1 & 0.2 & 0.7 & 0.9 \\
1 & 1 & 0.4 & 0.6 \\
0 & 0 & 0.5 & 0.4 \\
0.9 & 1 & 0.8 & 0.2
\end{bmatrix}
\end{array}
$$

If for a disease a particular symptom is prominent then degree of membership is high.

Now we want to find fuzzy relation that can correlate different type of plants to the symptoms:-

$R \circ S \rightarrow$ · max · min composition
max · product "
max dot "

This composition can be defined as:-

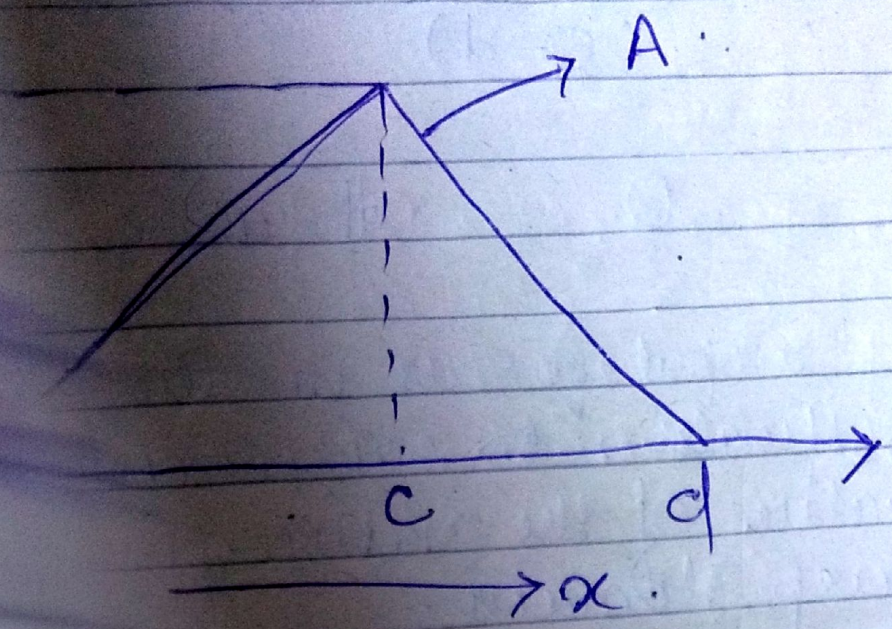
(Max min composition) $\mu_{R \circ S}(x, z) = \max_{y \in Y} \{ \mu_R(x, y) \wedge \mu_S(y, z) \}$

(max dot composition) $\mu_{R \circ S}(x, z) = \max_{y \in Y} \{ \mu_R(x, y) \cdot \mu_S(y, z) \}$

$\mu_{R \cdot S}(x, z) = \max \{ \mu_R(x, y) \wedge \mu_S(y, z) \}$

$P_1 S_1 = (0.6, 0.1)$; $P_1 S_1 = (0.6, 1)$;

Triangular



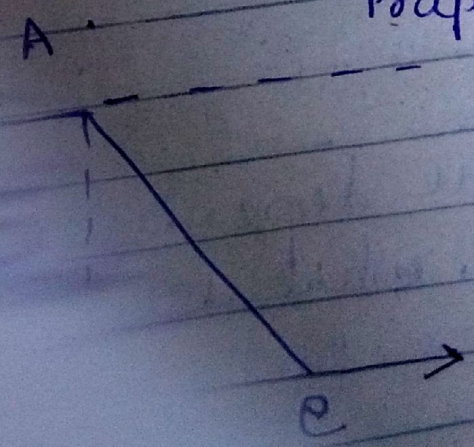
$\frac{(b-x)}{(b-c)}$ for $b \le x \le c$

$\frac{(d-x)}{(d-c)}$ for $c \le x \le d$

elsewise

value of a = 1.

Traperoidal



$b \le x \le c$

$c \le x \le d$

Q A, B are fuzzy set under universe of discourse $X = \{a, b, c, d, e\}$ and

$A = \{0.2/a, 0.6/b, 0.5/c, 1/d, 0.3/e\}$

$B = \{0.3/a, 0.7/b, 0.8/c, 0/d, 0.8/e\}$

$\bar{A} = \{0.8/a, 0.4/b, 0.5/c, 0/d, 0.7/e\}$

$\bar{B} = \{0.7/a, 0.3/b, 0.2/c, 1/d, 0.2/e\}$.

Complement of A wrt B.

$B/A = \max\{0, \mu_B(x) - \mu_A(x)\}$

$= [\max(0, 0.1/a), \max(0, 0.1/b), \max(0, 0.3/c), \max(0, -1/d), \max(0, 0.5/e)]$.

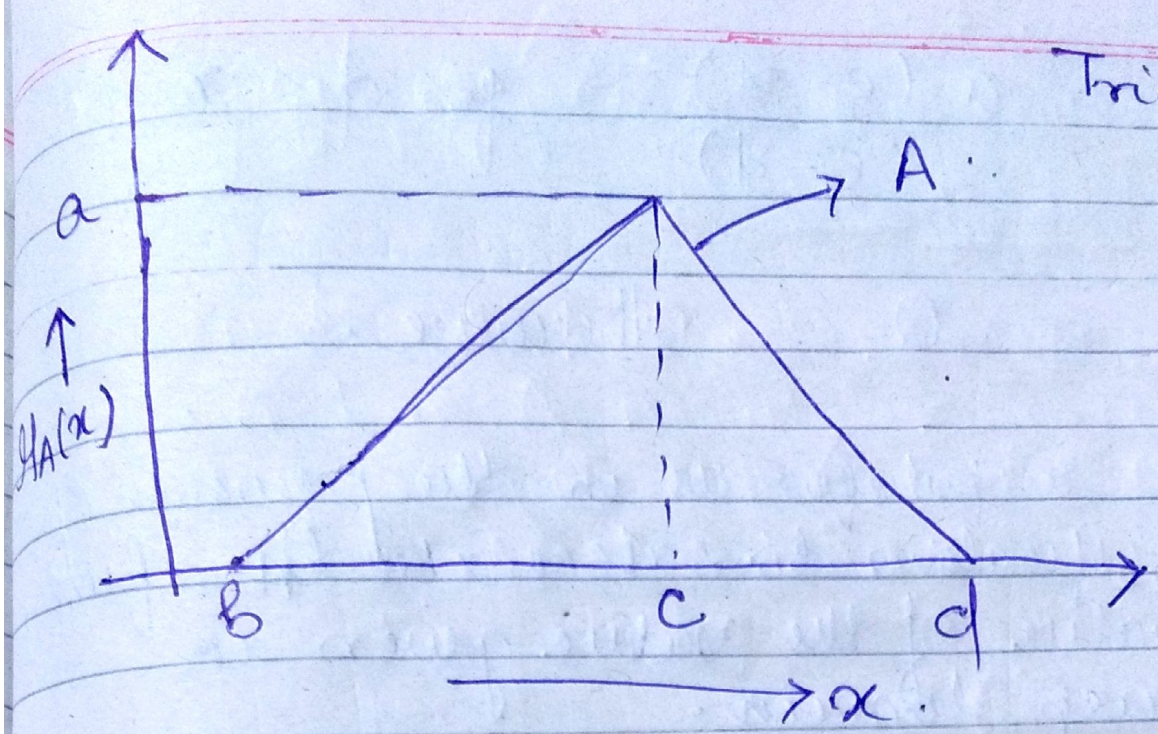
$= (0.1/a, 0.1/b, 0.3/c, 0/d, 0.5/e)$.

—x—

9/10/18.

Mathematical formulation of fuzzy sets.

Triangular



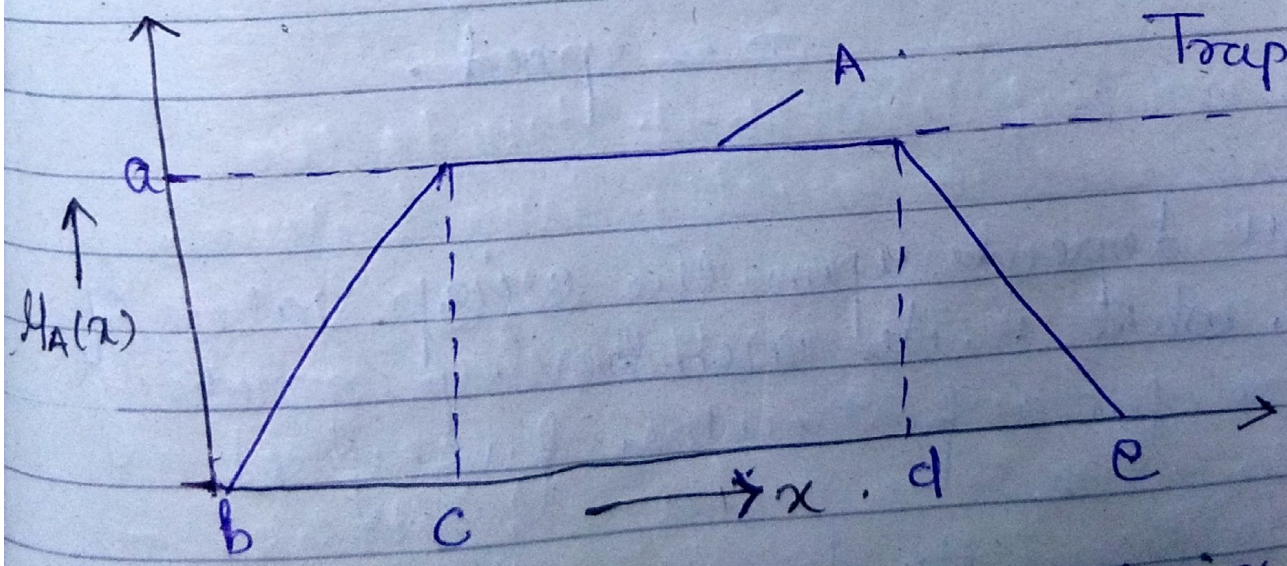
$$\mu_A(x) = \frac{a(b-x)}{(b-c)} \quad \text{for } b \le x \le c$$

$$\mu_A(x) = a\,\frac{(d-x)}{(d-c)} \quad \text{for } c \le x \le d$$

$$= 0 \quad \text{, otherwise}$$

Normally, the value of $a = 1$.

Trapezoidal



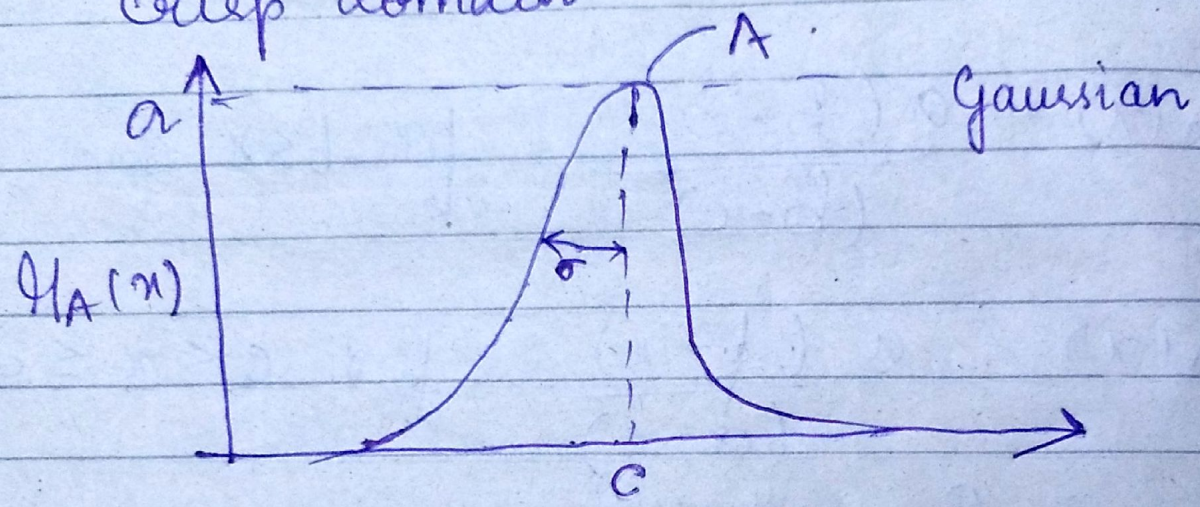
$$\mu_A(x) = a \cdot \frac{(b-x)}{(b-c)} \quad \text{for } b \le x \le c$$

for $e \le x \le d$

$$= \frac{a(e-x)}{(e-d)} \quad \text{for } d \le x \le e$$

$$= 0, \quad \text{Otherwise.}$$

This is useful for us in the process of fuzzification to calculate the fuzzy value of the value given in crisp domain.



$$\mu_A(x) = e^{-\frac{(x-c)^2}{2\sigma^2}}$$

σ = speed.

It all depends upon the crisp value of x, what is its membership grade of x for a particular fuzzy set.

<u>How to define the degree of fuziness of a fuzzy set.</u>

Now if μ is close to 0, the degree of fuzziness is uncertain or on the lower side or smaller and same for μ close to 1.

If its μ is close to 0.5 then degree of fuzziness is on higher side.

or, if the complement of a fuzzy set is similar to the original fuzzy set, then degree of fuzziness is on higher side.

<u>De Luca and Termini</u>.

Criterion to quantify the degree of fuzziness.

(1) $FUZ(F) = 0$ if $\mu_F(x) = \{0, 1\}$ $\forall x \in X$

↳ degree of fuzziness of set F.

(2) $FUZ(F) =$ Maximum if $\mu_F(x) = 0.5$ $\forall x \in X$.

(3) $FUZ(F^*) \leq FUZ(F)$ if F^* is the sharpened version of F.

Sharpened version means

(a) $\mu_{F^*}(x) \geq \mu_F(x)$ for $\mu_F(x) \geq 0.5$

(b) $\mu_{F^*}(x) \leq \mu_F(x)$ for $\mu_F(x) \leq 0.5$.

In the upper half, the membership grade of F^* is close to 1 and in the lower half

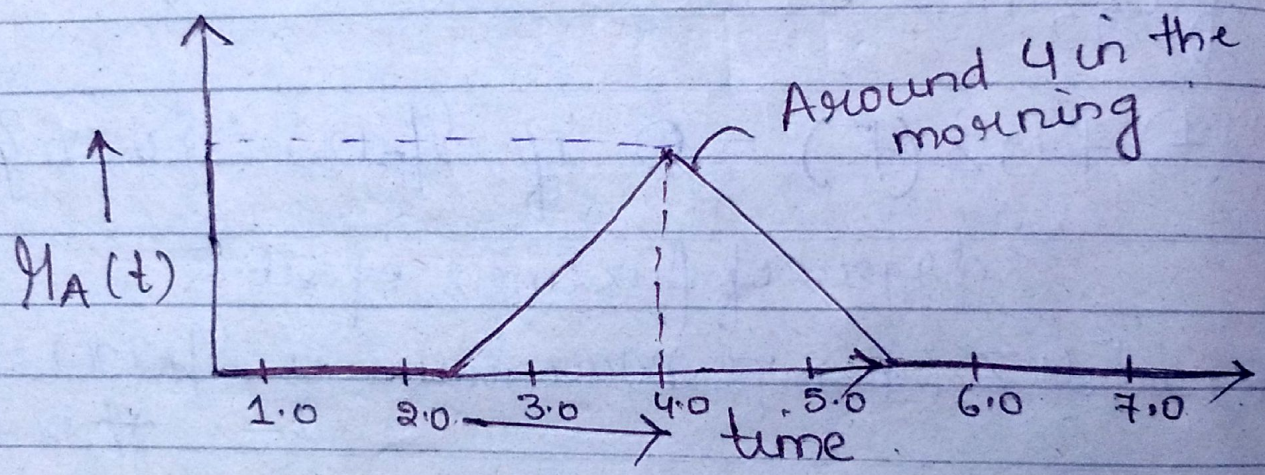
it is more closer to 0.
ie, membership grade of F* is always less than F.

How to evaluate the degree of fuzziness of a fuzzy set.

Suppose we are required to calculate the distance between A and $\bar{A}$, if it is small then degree of fuzziness is on higher side and vice-versa.

10/10/18 (Along with lat).

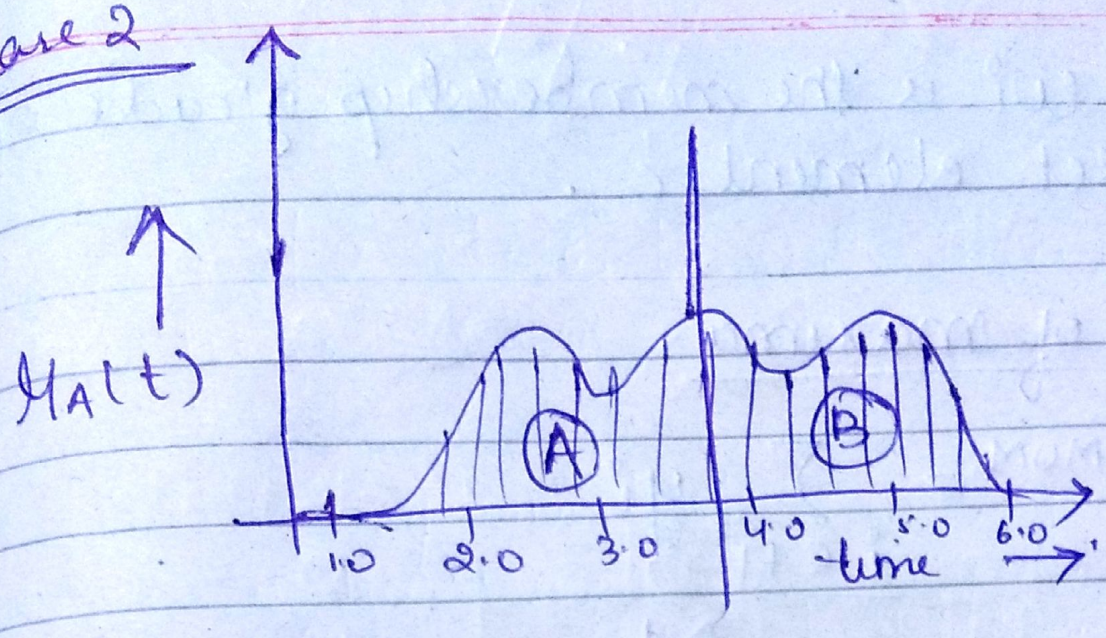
Defuzzification :-



The vertical line that divide the fuzzy set in exactly two halves, gives its crisp value.
so here it is 4 in the morning.

Now, if we have a fuzzy set like :-

Case 2



$A = B$.

Here, we will divide the area under the curve in two equal area part. This gives its crisp value.
For this, we have to use integration method which is time consuming & ~~one~~ computational burden increases.
For Defuzzification we use various methods like:-
(1) COA (Centre of Area) or Centroid method or Centre of Gravity.
(2) Mean of maxima.
(3) method of heights.

~~Geometric~~
For finite Universe of discourse,

① COA

$$y^{COA} = \left(\sum_{i=1}^{n} w_i * y_i\right) \Big/ \left(\sum_{i=1}^{n} w_i\right)$$

where w_i is the membership grade of finite set element i.

② Mean of maxima

$$y^{MOM} = \frac{\sum_{i \in H} y_i}{M}$$

H is a set comprising of all those values where membership grade is maximum.

M is the cardinality of H.

③ Method of Heights

$$y^{MOH} = \frac{\sum_{i \in G} w_i y_i}{\sum_{i \in G} w_i}$$

G is the alpha cut of a fuzzy set where values are y_i.
G is a crisp set.

y_i are that values whose value is more than or equal to α.

More generalised method is MOH, by setting valu