

06/09/19.

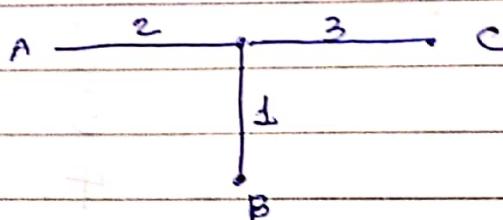
## Molecular Evolution

Distance Matrix

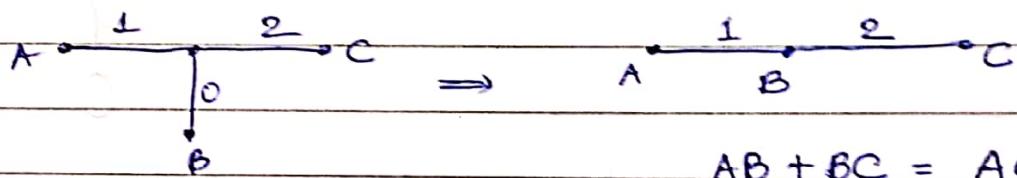
Dij

Reconstruction

Distance based

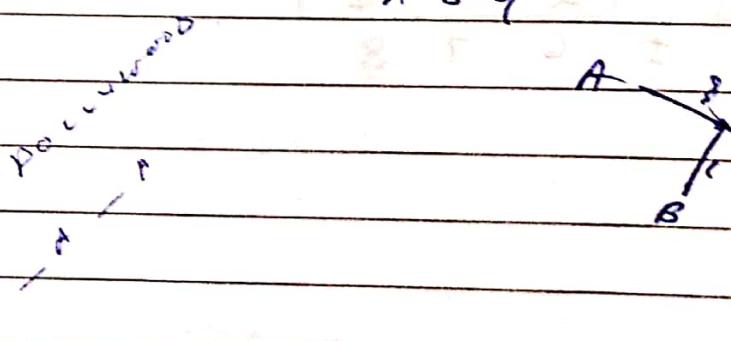
phylogeny  
algorithmDegenerate  
Triples.Shortening  
hanging edge

Reduce all leaves by 1.



$$AB + BC = AC$$

After sufficient number of steps, we will get such degenerate triples.



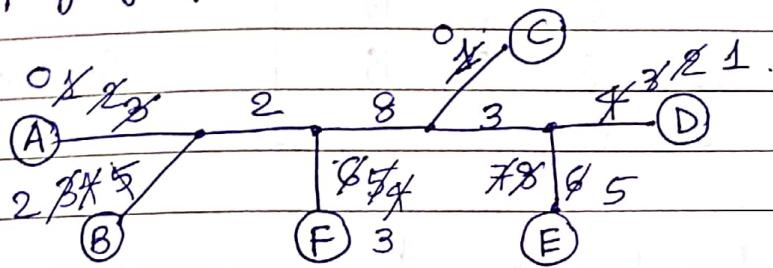
11/09/19

Page No.:

Date:

Yours

given a distance matrix, can we recreate a phylogeny tree?



	A	B	C	D	E	F	
A	0	8	14	20	24	11	
B	8	0	16	22	26	13	
C	14	16	0	8	12	15	
D	20	22	8	0	12	21	
E	24	26	12	12	0	25	
F	11	13	15	21	25	0	

distance b/w  
leaves - reduce

reduce 1 from  
each branch  
i.e. 2 from  
each distance

0	6	12	18	22	9
6	0	14	20	24	11
12	14	0	6	10	13
18	20	6	0	10	19
22	24	10	10	0	23
9	11	13	19	23	0

New, C is  
degenerate.  
 $AC + CD = AD$ .  
 $\therefore C$  is  
between A and  
D.

$\Rightarrow$  remove C.

	A	B	D	E	F
A	0	4	16	20	7
B	4	0	20	24	11
D	16	20	0	10	19
E	20	24	10	0	21
F	7	11	19	21	0

Nothing is  
degenerate.

	A	B	D	E	F	
A	0	2	14	18	5	$BA + AD = BD$
B	2	0	18	22	9	A is degenerate.
D	14	18	0	8	17	Remove A.
E	18	22	8	0	19	
F	5	9	17	19	0	

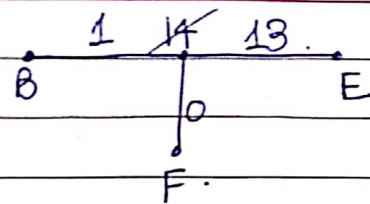
	B	D	E	F	
B	0	14	18	5	$BD + DE = BE$
D	14	0	4	13	$\therefore D$ is degenerate.
E	18	4	0	17	
F	5	13	17	0	

	B	E	F	
B	0	16	3	Reduce further.
E	16	0	15	
F	3	15	0	

	B	E	F	
B	0	14	1	$EF + FB \rightarrow BE$
E	14	0	13	
F	1	13	0	$\therefore F$ is degenerate.

	B	E	
B	0	14	
E	14	0	

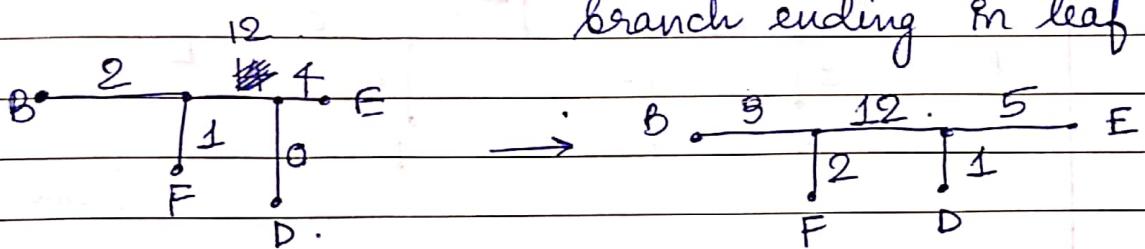
Now, we go reverse to build tree from BE.



Now, we insert F in between — give 0 to its branch.

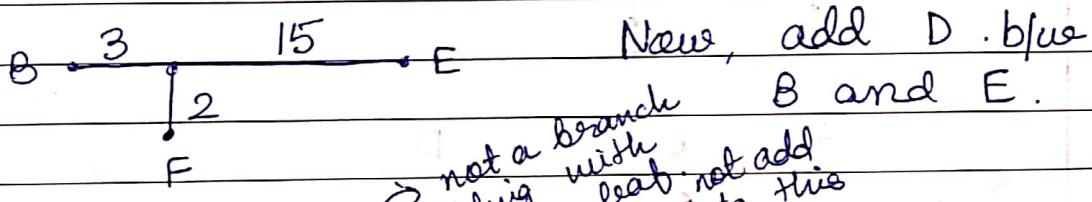
Write BF distance and FE distance.

Now, increase 1 in each branch ending in leaf.

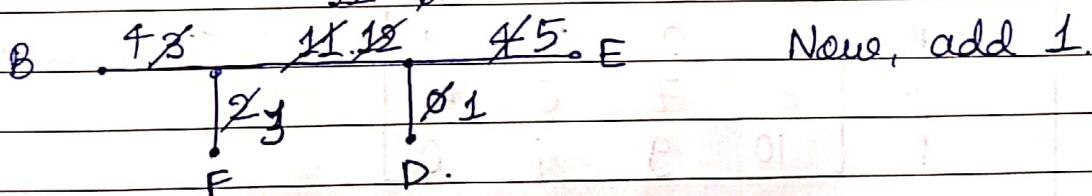


We had reduced 2 times.

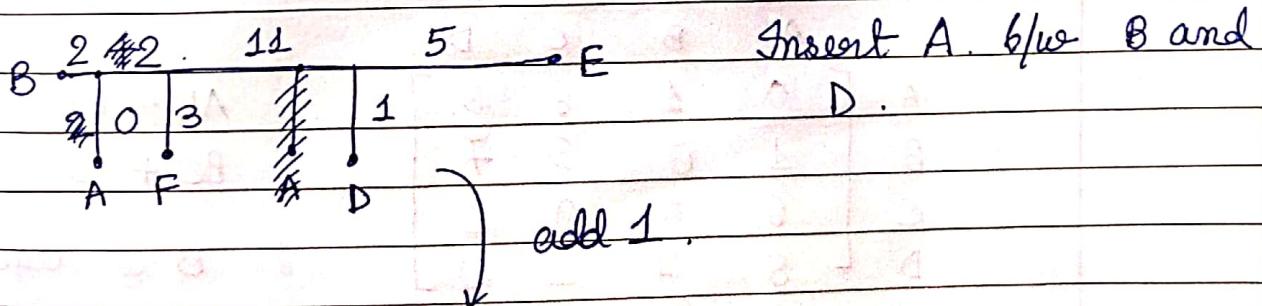
NO.



Now, add D b/w B and E.

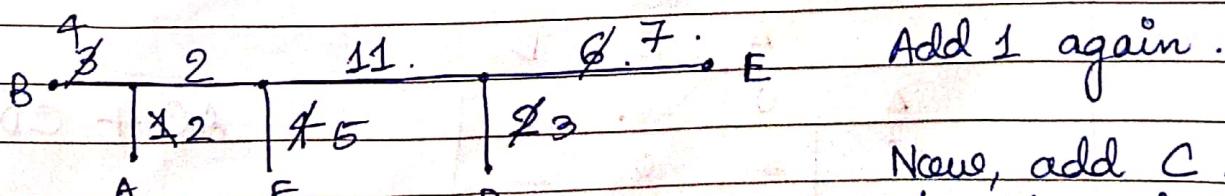


Now, add 1.



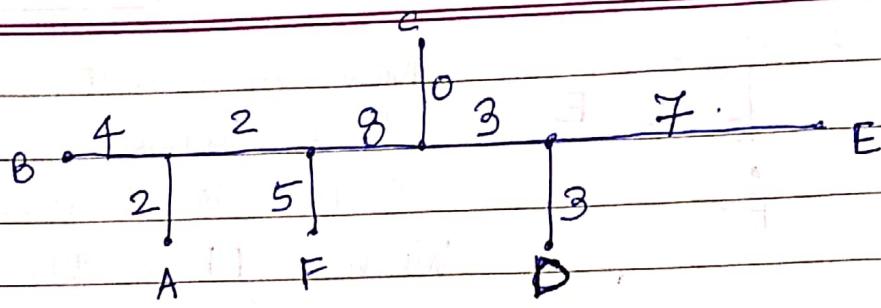
Insert A. b/w B and D.

add 1.

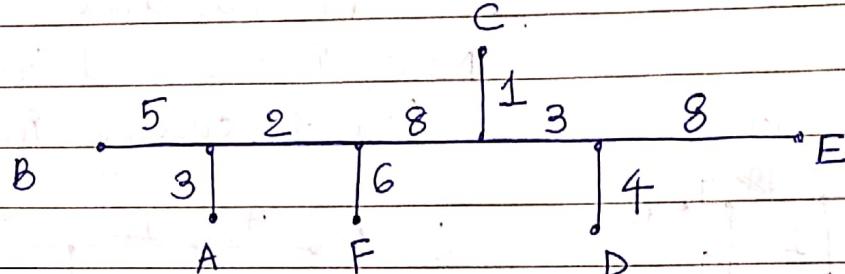


Add 1 again.

Now, add C. b/w A and D.

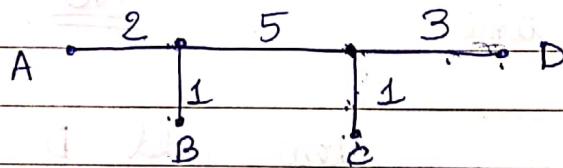


Add 1  
to each  
branch



Final tree.

Ex)



	A	B	C	D
A	0	3	8	10
B	3	0	7	9
C	8	7	0	4
D	10	9	4	0

Decrease by 2.

	A	B	C	D
A	0	1	6	8
B	1	0	5	7
C	6	5	0	2
D	8	7	2	0

$$AB + BD \rightarrow AD.$$

$$BC + CD \rightarrow B$$

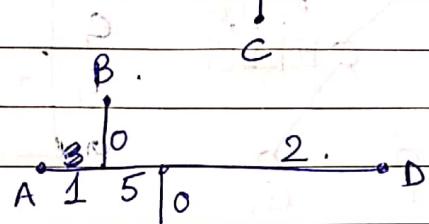
∴ B is degenerate

	A	C	D
A	0	6	8
C	6	0	2
D	8	2	0

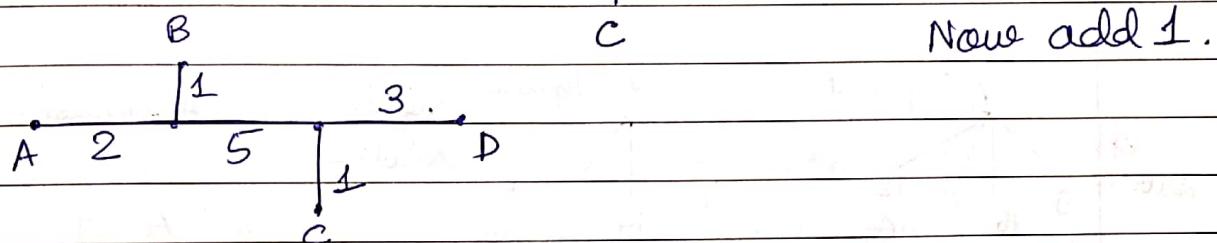
$$AC + CD = AD$$

∴ C is degenerate

$$\begin{array}{cc} A & D \\ \left[ \begin{array}{cc} 0 & 8 \\ 8 & 0 \end{array} \right] & \end{array} \quad \begin{array}{c} 6 \\ | \\ 0 \\ | \\ 2 \end{array} \quad \text{Add } C$$



Add B.



Now add 1.

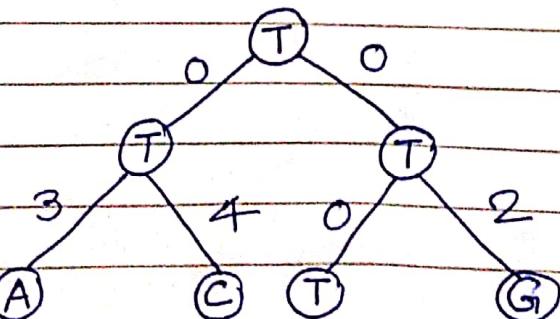
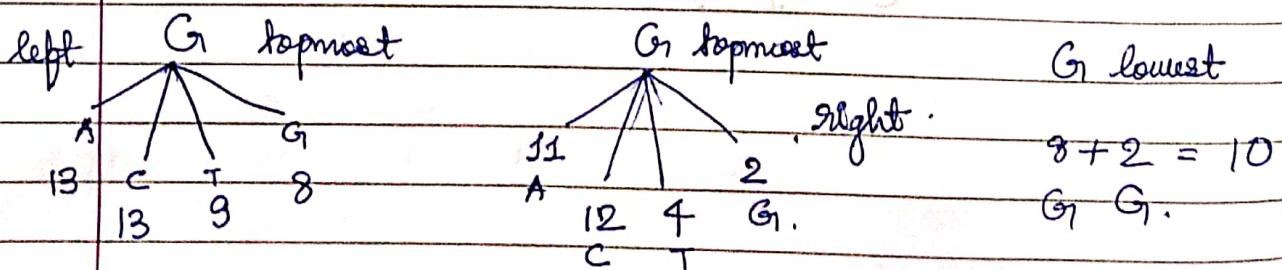
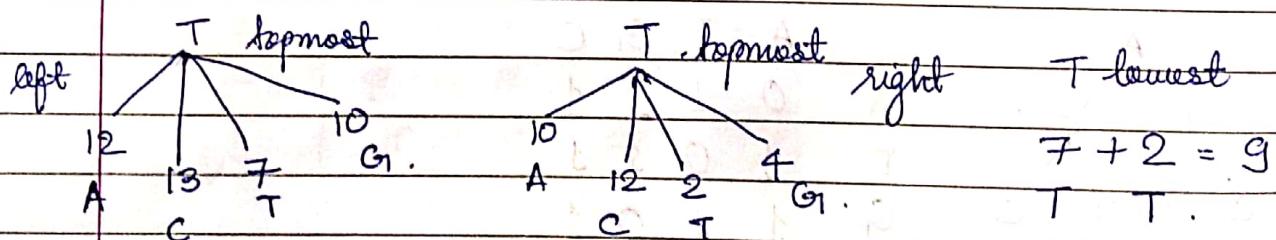
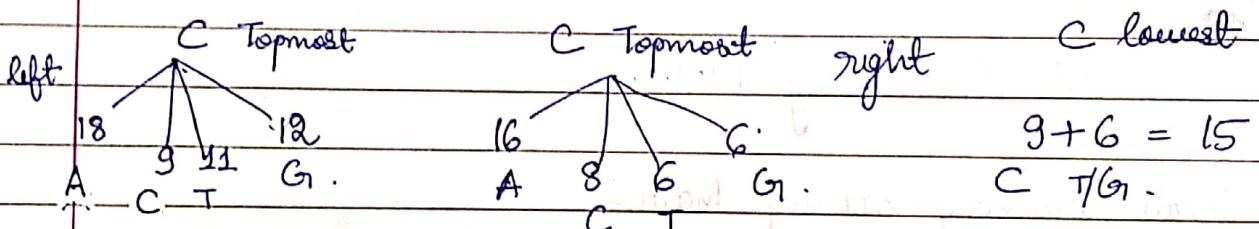
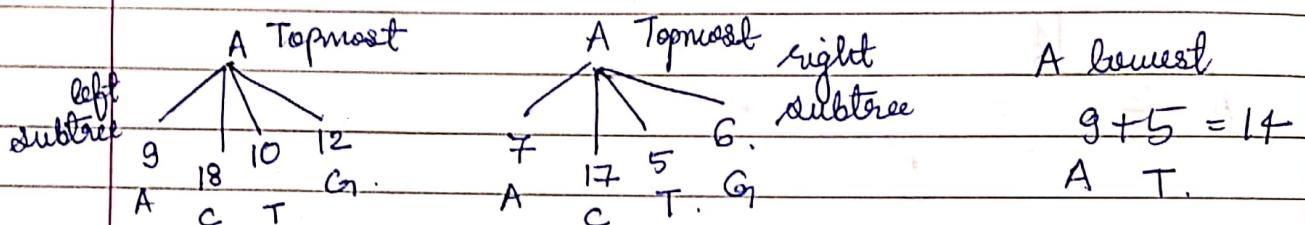
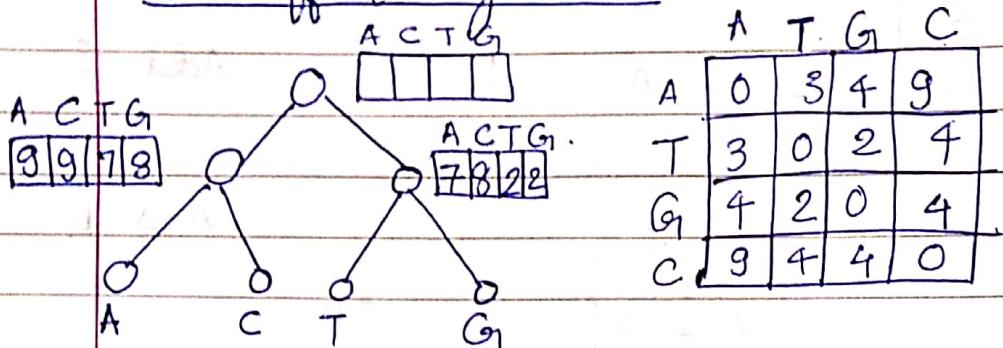
13/09/19

### Small Parsimony Matrix

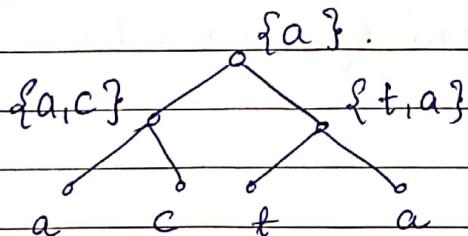
### Small Parsimony Scoring Matrix

A	T	G	C
A	0	1	1
T	1	0	1
G	1	1	0
C	1	1	0

## Sankoff's Algorithm



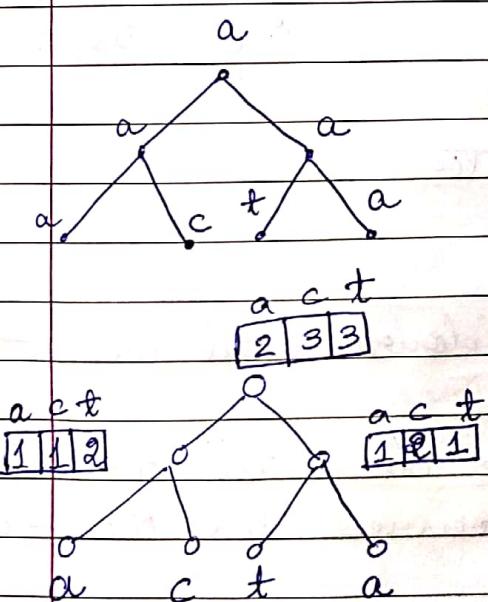
## Fitch's Algorithm



make labels.  
go from bottom to top,  
from top, start  
labelling.

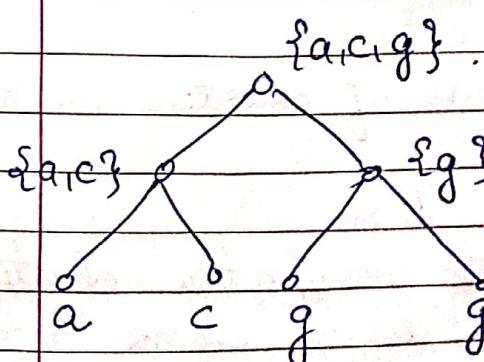
While labelling

- 1) single element
    - take that
  - 2) multiple element
    - if parent element in that set, take parent
    - else pick randomly any one of them.
- if  $\cap$  is non zero, take  $\cap$   
if  $\cap$  is zero, take  $\cup$ .



small parsimony.

a	c	t	
a	0	1	1
c	1	0	1
t	1	1	0



Fitch's Algorithm and  
Sankoff's Algorithm gives  
same answers iff the  
weights are 1.

Evolutionary tree structure  
is given - we need to  
label the intermediate  
nodes such that path  
is minimum.

## Large Parsimony Problem.

Strings at leaves instead of characters.

18/09/19

## Clustering

nodes are not linear  
can be 3-D or more  
bottom up.

some measure to group  
converge to one node.

→ start with  
Distance Matrix  
→ two closest -  
group using some  
function.

each group is called a cluster.

From top to bottom,

- 1) start from top, go down until we get no. of groups.
- 2) set a threshold, go down till threshold and no. of groups formed till then - are

## K-means clustering

no. of  
clusters  
needed.

start from 'k' → define no. of clusters we want.  
define centres.

compare all points with them.

⇒ group the closer ones to respective clusters.  
⇒ find centre of clusters using some averaging method.

⇒ Repeat process.

- Variations :
- restrict the movement of centre so that there are no major changes and it stabilizes.
  - take some existing point as centre instead of random centres in the beginning.

Ex) K-means clustering

Distance Matrix - Euclidean distance. (3 clusters)

	1	2	3	4	5	6	7	8	10
1	0	$\sqrt{25}$	$\sqrt{36}$	$\sqrt{13}$	$\sqrt{50}$	$\sqrt{52}$	$\sqrt{65}$	$\sqrt{5}$	
2		0	$\sqrt{37}$	$\sqrt{18}$	$\sqrt{25}$	$\sqrt{17}$	$\sqrt{10}$	$\sqrt{20}$	
3			$\sqrt{0}$	$\sqrt{25}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{53}$	$\sqrt{4}$	
4				0	$\sqrt{13}$	$\sqrt{17}$	$\sqrt{52}$	$\sqrt{2}$	
5					0	$\sqrt{2}$	$\sqrt{45}$	$\sqrt{25}$	
6						0	$\sqrt{29}$	$\sqrt{29}$	
7							0	$\sqrt{58}$	
8								0	
9	A1 (2,10)	A3 (8,4)	A5 (7,5)	A7 (1,2)					
10	A2 (2,5)	A4 (5,8)	A6 (6,4)	A8 (4,9)					

Let A1, A4, A7 be those centres.  
 $\downarrow$        $\downarrow$        $\downarrow$   
I      II      III.

For A2, A3, A5, A6, A8 - find distance from the three centres and cluster them in one which is closest.

new clusters formed

$$1 : \{A1\} \quad 3 : \{A2, A7\}$$

$$2 : \{A3, A4, A5, A6, A8\}$$

Find centres of new clusters

$$= (\text{avg of } x, \text{ avg of } y)$$

$$\begin{aligned} \text{new centres} &= (2, 10) \quad C_I \\ &= (6, 6) \quad C_{II} \\ &= (1.5, 3.5) \quad C_{III} \end{aligned}$$

$$A_1 : C_I$$

$$A_2 : (2, 5). \quad C_I = \sqrt{25} \quad C_{III}$$

$$C_{II} = \sqrt{10}$$

$$C_{III} = \sqrt{25} \quad \checkmark$$

$$A_3 : (8, 4) \quad C_I = \sqrt{72} \quad C_{II}$$

$$C_{II} = \sqrt{8} \quad \checkmark$$

$$C_{III} = \sqrt{(6.5)^2 + (0.5)^2}$$

$$A_4 : (5, 8) \quad C_I = \sqrt{13} \quad C_{II}$$

$$C_{II} = \sqrt{5} \quad \checkmark$$

$$C_{III} = \sqrt{3.5^2 + 4.5^2}$$

$$A_5 : (7, 5) \quad C_I = \sqrt{50}$$

$$C_{II} = \sqrt{2} \quad \checkmark$$

$$C_{III} = \sqrt{5.5^2 + 1.5^2}$$

$$A_6 : (6, 4) \quad C_I = \sqrt{52}$$

$$C_{II} = \sqrt{4} \quad \checkmark$$

$$C_{III} = \sqrt{4.5^2 + 0.5^2}$$

$$A_7 : (1, 2) \quad C_I = \sqrt{65}$$

$$C_{II} = \sqrt{41}$$

$$C_{III} = \sqrt{(0.5)^2 + (1.5)^2}$$

$$A_8 : (4, 9) \quad C_I = \sqrt{5} \quad \checkmark$$

$$C_{II} = \sqrt{13}$$

$$C_{III} = \sqrt{2.5^2 + 5.5^2}$$

A

$$C_I : \{ A_1, A_8 \} \quad \text{centres} \quad (3, 9.5)$$

$$C_{II} : \{ A_3, A_4, A_5, A_6 \} \quad (6.5, 5.25)$$

$$C_{III} : \{ A_2, A_7 \} \quad (1.5, 3.5)$$

We will stop iterating when no changes are seen in the clusters.

Multiple runs to decide the seed centres.

Unsupervised learning algorithm.

### Eg) Hierarchical Clustering

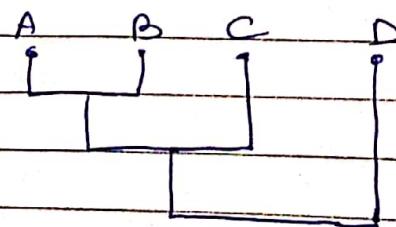
single link : distance b/w two clusters is the shortest distance b/w a pair of elements from two clusters.

complete link : longest distance

	A	B	C	D	start with all points as clusters.
A	0	1	4	5	
B		0	2	6	
C			0	3	A-B (shortest).
D				0	AB - clustered.

	AB	C	D	now, we merge AB and C (shortest)
AB	0	2	5	
C		0	3	
D			0	

	ABC	D	now merge ABCD.
ABC	00	03.	
D	03	00	



Dendrogram

Dendrogram  
for single  
link

Complete linkage.

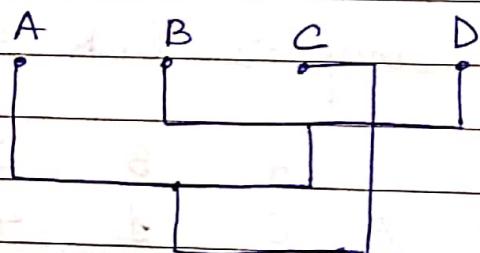
	A	B	C	D
A	0	1	4	5
B		0	2	6
C			0	3
D				0

Merge B and D

	A	C	BD
A	0	4	5
C		0	(3)
BD			0

Now merge A and BD

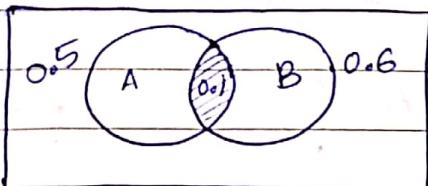
	ABD	C
ABD	0	4
C	4	0



24/9/19.

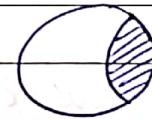
Page No.:	
Date:	youva

## Baye's Theorem



$$\begin{aligned}P(A) &= 0.5 \\ P(B) &= 0.6 \\ P(A \cap B) &= 0.1\end{aligned}$$

$$\begin{aligned}P(B|A) &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{0.1}{0.5}\end{aligned}$$



$P(A) \rightarrow$   
now becomes  
1.

$$\Rightarrow P(A \cap B) = P(A) * P(B|A)$$

Similarly  $P(A \cap B) = P(B) * P(A|B)$

Since  $P(A|B) = \frac{P(A \cap B)}{P(B)}$

$$\begin{aligned}\therefore P(B) * P(A|B) &= P(A) * P(B|A) \\ \Rightarrow P(B|A) &= \frac{P(A|B) * P(B)}{P(A)}\end{aligned}$$

Reverse the causal probabilities

Cause then effect.

But with Baye's theorem, we can find cause using effect

Question :  $P(\text{Blue}) = 0.15$

$$P(\text{Green}) = 0.85$$

$$P(\text{truth}) = 0.8$$

$$P(\text{failed to identify}) = 0.2$$

$P(\text{Blue})$  and  $P(\text{truth})$ .

85/ 151.

80%	68	12
	A	B

20%	C 17	3
		D

Universe - all cars identified as blue.

A : green identified as green

B : blue identified as blue.

$P(\text{car was blue})$

$$= \frac{12}{17+18}$$

C : green identified as blue.

D : blue identified as green.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{12}{29}$$

$$= \frac{12}{29}$$

$$P(A|B) \approx 40\%$$

Q.  $P(\text{Cancer}) = 0.001$ .

$P(\text{accurate}) = 0.9$

	①	②	
90%	8991	9.	10,000 people
10%	③ 999	④ 1	$\frac{9990}{9990} = 8991$

— NO.

Ans?

Test	+ve	-ve
+ve	98%	10% False +ve
-ve	2%	90% False -ve

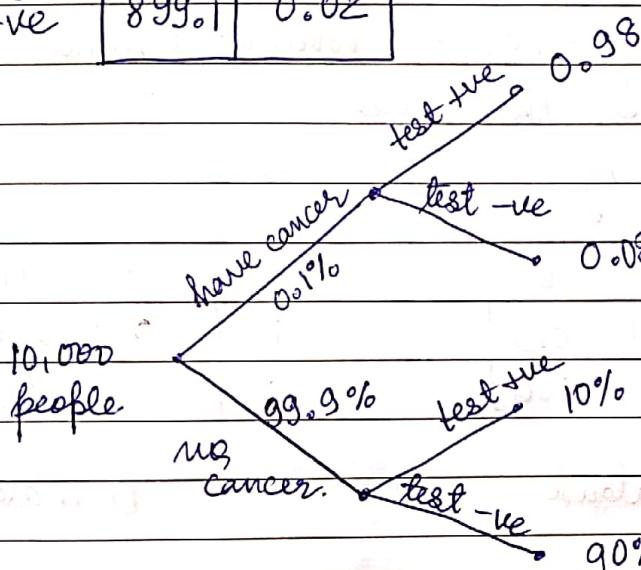
No  
cancer      Cancer

Test +ve	99.9	0.98
Test -ve	0.001	0.02

$$P(\text{Cancer} | \text{test +ve})$$

$$= \frac{0.98}{0.98 + 0.02}$$

$$= 0.98$$



Q.

universe - 100 people.

30% reliable - guilty      10% - not reliable.

98% reliable - innocent.

Of 100 suspects, only 12 are actually involved.

guilty	90/100 * 12	10/100 * 12
innocent	0.02 * 98	0.98 * 88

$$\text{total} = 12$$

test +ve - guilty

test -ve - innocent

$$\text{total} = 88 \quad P(\text{guilty} | \text{test is +ve})$$

$$= \frac{108/10}{108/10 + 17/100}$$

$$= \frac{108/10}{108/10 + 17/100}$$

$$= \frac{108/10}{108/10 + 17/100}$$

25/9/19



Page No.:  
Date:  
Youshika

"If you feel stressed, you should do meditation //  
(Words of Kushekar Sir)  
— didhi (ACMBai)

HMM (Bayesian NW - reversal of effect and cause)

Take the rings (of trees' trunks).

Measure thickness of rings - which depends on temperature over the year.

One year - one ring

	H	C		S	M	L
H	0.7	0.3	H	0.1	0.4	0.5
C	0.4	0.6	C	0.7	0.2	0.1

→

A mature

B mature

))))))))))))))) → 100 years

Count hot years and cold years

	H	C		last 100 years
π	0.6	0.4	60 years - Hot	
			40 years - Cold	

π mature

\* HMM will look like the following :

gives the most probable state sequence  
given the observation

- HMM machine

⇒ Understand cause given effect

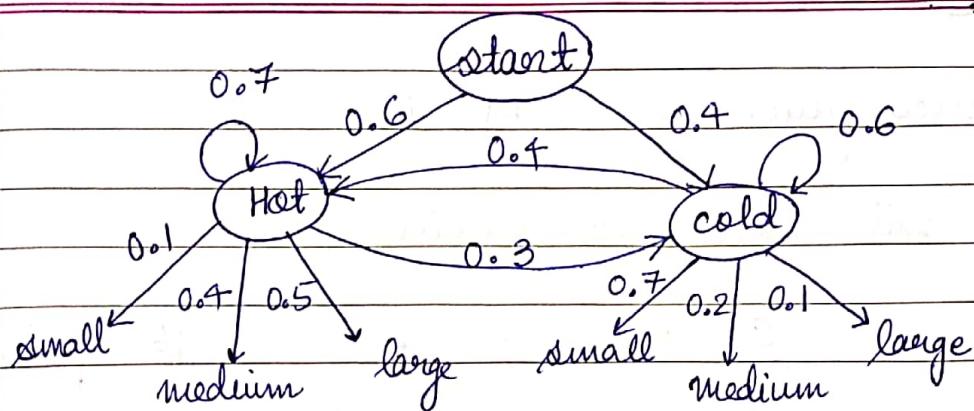
⇒ also make prediction - once machine is made, we can run it for future as well

## HMM Machine.

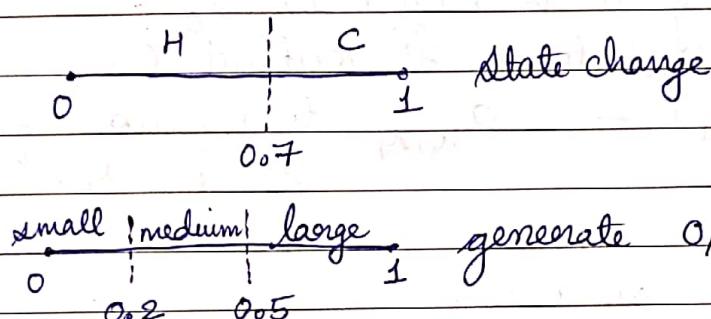
Page No.:

Date:

XOUVX



Generate a random no. between 0 and 1.



→ Given a sequence of data, we can generate the HMM machine as well, given no. of states.

Doesn't require prior knowledge.

Unsupervised learning

Learning of given data set to infer what the states are.

27/09/9.

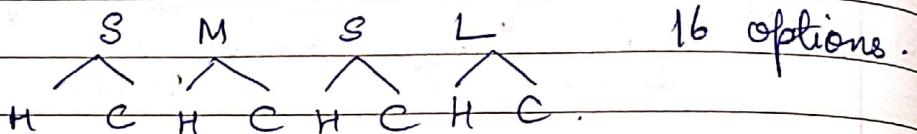
Page No.:

Date:

Yours

Sequence given : S M S L .

Each will have two choices .

 $\pi_{x_0}$  — using  $\pi$  to start in  $x_0$ . $b_{x_0}(0_0)$  — state  $x_0$  at output  $0_0$ . $a_{x_0, x_1} b_{x_1}(0_1)$  —  $a_{x_0 x_1}$  transition from  $x_0$  to  $x_1$ .
 $(0, 1, 0, 2)$  — sequence.  
 ↓      ↓      ↓  
 S      M      L .

$$P(HHCC) = (0.6)(0.1)(0.7)(0.4)(0.3)(0.7)(0.6)(0.1)$$

↓      ↓      ↓      ↓      ↓      ↓      ↓  
 P(H) . H  $\rightarrow$  S    H  $\rightarrow$  H .    H  $\rightarrow$  M .    H  $\rightarrow$  C    C  $\rightarrow$  S    C  $\rightarrow$  C  
 O/P transition      O/P transition      O/P

Now in this table [ state probability normalised (divide by) ]  
 16 states       $P_i^o$        $\sum P_i^o$

Sequence required sum first 8 elements having first H in order.  
 Sum (H).

Similarly sum (C).  
 In the same way we get

Sum (HH)  
 Sum (HC)  
 Sum (CH)  
 Sum (CC). } not needed.

Sum (1-H)

↓  
 sum of probabilities for  
 states where 1<sup>st</sup> element is H.

1 0 1 2 3,

element

0      1      2      3

$P(H)$  sum (0-H) sum (~~1-H~~) sum (2-H) sum (3-H)  
 $P(C)$  sum (0-C) sum (1-C) sum (2-C) sum (3-C)

These sums are additions of normalised probability  $\therefore$  columns 0, 1, 2, 3 - add to 1 each.

Take the best of 2 for each column  
 → state most probable.

first method : CCHH (Dynamic Programming)  
 second method : CHCH (HMM)

Machine -  $\lambda$   
 observation sequence - O       $P(O|\lambda) = \sum_{\text{states}} P(O, x|\lambda)$

If we have two states when we start reading DNA  
 gene      not a gene.  
 (hot)      (cold)

Problem 2 How large will the table be?

If observation sequence is 20 chars, table becomes  
 $2^{20} \Rightarrow$  Time complexity increases  
 Space complexity increases.

exponential

$\lambda = (A, B, \pi)$ , O  
 find  $P(O|\lambda)$ .

$\lambda = (A, B, \pi)$  - A model.

$O = (O_0, O_1, O_2 \dots)$  is a series of observations

Find  $P(O|\lambda)$ .

Let  $X = (x_0, x_1 \dots x_{T-1})$  be state sequence.

$$P(O|X, \lambda) = b_{x_0}(O_0) b_{x_1}(O_1) \dots b_{x_{T-1}}(O_{T-1}).$$

$$P(X|\lambda) = \pi_{x_0} a_{x_0, x_1} a_{x_1, x_2} \dots a_{x_{T-2}, x_{T-1}}$$

$$\underline{P(O|X, \lambda)} \quad P(O, X|\lambda) = \frac{P(O \cap X \cap \lambda)}{P(\lambda)}$$

$$\begin{aligned} P(O, X|\lambda) \cdot P(X|\lambda) &= \frac{P(O \cap X \cap \lambda)}{P(X \cap \lambda)} \cdot \frac{P(X \cap \lambda)}{P(\lambda)} \\ &= P(O|X, \lambda) \cdot P(X|\lambda). \end{aligned}$$

$$P(O, X|\lambda) = P(O|X, \lambda) \cdot P(X|\lambda) \quad \text{one row in the table} \\ \text{given } X.$$

For all possible  $X$ 's (states)

$$P(O|\lambda) = \sum_X P(O, X|\lambda)$$

$$= \sum_X P(O|X, \lambda) \cdot P(X|\lambda)$$

Forward Algorithm ( $\alpha$ -pass).

$$\alpha_t(i) = P(O_0, O_1, \dots, O_t, x_t = q_i | \lambda)$$

seen  $t$  states

in state  $q_i$  after seeing  $t$  states

no. of states = # of cols.

$\alpha_0(H)$	$\alpha_0(C)$
$\alpha_1(H)$	$\alpha_1(C)$
:	:
$\alpha_{T-1}(H)$	$\alpha_{T-1}(C)$

We store this table.

# of rows = # of ~~ops in sequence~~ <sup>ops in</sup> possible

1) Let  $\alpha_0(i) = \pi_i b_i(O_0)$  for  $i = 0, 1, \dots, N-1$ .

2) For  $t = 1, 2, \dots, T-1$  and  $i = 0, 1, \dots, N-1$

$$\alpha_t(i) = \left[ \sum_{j=0}^{N-1} \alpha_{t-1}(j) a_{ji} \right] b_i(O_t).$$

3) Then,

$$P(O|A) = \sum_{i=0}^{N-1} \alpha_{T-1}(i).$$

	H	C	
S	$0.6 * 0.1 = 0.06$	$0.4 * 0.7 = 0.28$	sum (last row)
M	0.0616	0.0372	
S	0.0058	0.02856	
L	0.00774	0.00168	

= 0.0096.

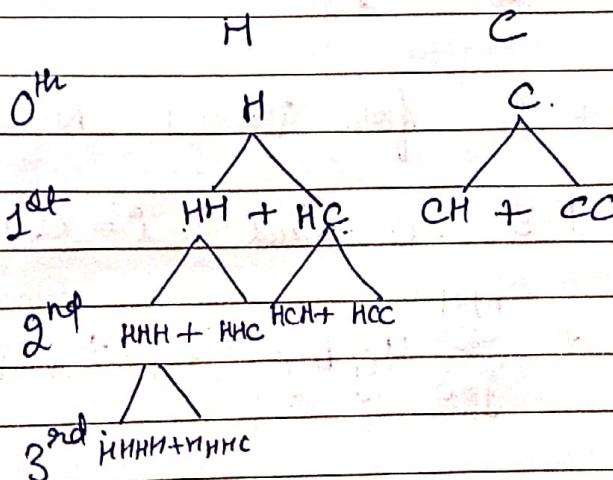
↓

$P(O|A)$ .

For M. [initial \* transition  
\* O/P ↓ (sum for H and C) - same for C.]

↓

sum of middle column.



no. of multiplications to calculate each value  
after first saw - constant.

Computation : exponential to polynomial

Table : m rows or columns  
 $O(mn)$ .

Preserve of matrix

Q. Given

$\beta$ -pass or backward algorithm.

Probabilities for final states = 1.

	H	C	
S	0.0302	0.02792	
M	0.0812	0.1244	
S	0.38	0.26	
L	1	1	

$\gamma^* = \alpha^* \beta$

1) Let  $\beta_{T-1}(i) = 1$ . for  $i = 0, 1, \dots, N-1$

2) For  $t = T-2, T-3, \dots, 0$  and  $i = 0, 1, \dots, N-1$   
compute

$$\beta_t(i) = \sum_{j=0}^{N-1} a_{ij} b_j (\beta_{t+1})^j$$

Tell my teacher about  
for next time  
seriously??  $\ominus$

Page No.	_____
Date	_____

$$\gamma_t(i) = P(x_t | q_i^o | O, \lambda)$$

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)}$$

HMM probabilities

most likely state at time  $t$  is the state  $q_i^o$  for which  $\gamma_t(i)$  is maximum where the maximum is taken over the index  $i$ .

$$\text{Digamma} \quad \gamma_t(i, j) = P(\alpha_t = q_i^o, \alpha_{t+1} = q_j^o | O, \lambda).$$

state  $i$  to  $j$  at time  $t$

$$1) \pi_{i^o} = \gamma_0(i^o)$$

$$2) a_{ij} = \sum_{t=0}^{T-2} \gamma_t(i, j) / \sum_{t=0}^{T-2} \gamma_t(i)$$

$$3) b_{j^o(k)} = \sum_{t \in \{0, 1, \dots, T-1\}} \gamma_t(j^o) / \sum_{t=0}^{T-1} \gamma_t(j^o)$$

$$= 1 - \sum_{t=0}^{T-1} \gamma_t(j^o) / \sum_{t=0}^{T-1} \gamma_t(j^o)$$

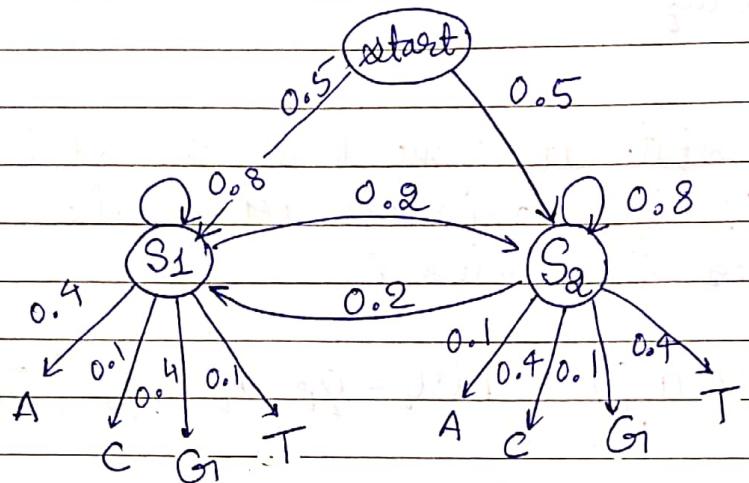
01/10/19

## HMM for DNA Sequence.

$$P(S_1 | S_1) = 0.8$$

$$P(S_2 | S_1) = 0.2$$

$$P(S_1 | S_2)$$



CGTCAG : given sequence.

$\alpha$ :	$S_1$	$S_2$
C	$0.5 * 0.1 = 0.05$	$0.5 * 0.4 = 0.2$
G	$0.05 * 0.8 * 0.4 +$ $0.5 * 0.2 * 0.4 = 0.032$	$0.2 * 0.8 * 0.1 +$ $0.5 * 0.2 * 0.1 = 0.017$
T	$0.032 * 0.8 * 0.1 +$ $0.017 * 0.8 * 0.1 = 0.0029$	$0.082 * 0.2 * 0.4 +$ $0.017 * 0.8 * 0.4 = 0.008$
C	$0.0029 * 0.8 * 0.1 +$ $0.008 * 0.2 * 0.1 = 0.000392$	$0.0029 * 0.2 * 0.4 +$ $0.008 * 0.8 * 0.4 = 0.0028$
A	$0.000392 * 0.8 * 0.4 +$ $0.0028 * 0.2 * 0.4 =$ $0.0003488 = 0.0003488$	$0.000392 * 0.2 * 0.1 +$ $0.0028 * 0.8 * 0.1 = 0.0002312$
G	$0.0003488 * 0.8 * 0.4 +$ $0.0002312 * 0.2 * 0.4$ $= 0.00013011$	$0.0003488 * 0.2 * 0.1 +$ $0.0002312 * 0.8 * 0.1$ $= 0.00025472$
Sum these.		

$$P(0|2) = 0.000155582 \\ = 1.55582 \times 10^{-4}$$

$\beta.$  :

	S1	S2
C	$0.002176 * 0.8 * 0.4 + 0.005056 * 0.8 * 0.1 + 0.002176 * 0.025056 * 0.2 * 0.1 = 0.00044744$	$0.005056 * 0.8 * 0.1 + 0.002176 * 0.025056 * 0.2 * 0.1 = 0.00057856$
G	$0.01216 * 0.8 * 0.1 + 0.01504 * 0.2 * 0.4 = 0.002176$	$0.01216 * 0.2 * 0.1 + 0.01504 * 0.8 * 0.4 = 0.005056$
T	$0.112 * 0.8 * 0.1 + 0.04 * 0.2 * 0.4 = 0.01216$	$0.112 * 0.2 * 0.1 + 0.04 * 0.8 * 0.4 = 0.01504$
C	$0.34 * 0.8 * 0.4 + 0.16 * 0.2 * 0.1 = 0.112$	$0.34 * 0.2 * 0.4 + 0.16 * 0.8 * 0.1 = 0.04$
A	$1 * 0.8 * 0.4 + 1 * 0.2 * 0.1 = 0.34$	$1 * 0.2 * 0.4 + 1 * 0.8 * 0.1 = 0.16$
G	1	1

$\gamma$  : multiply corresponding elements of  $\alpha$  and  $\beta$  sum of last row of  $\alpha$ ,  $\rightarrow P(0|2)$ .

sum of each row should sum to 1.

$\gamma$ :

	S1	S2
C	0.25627	0.74372
G	0.44455	0.55244
T		
C		
A		
G		