WRITTEN ASSIGNMENT - I
INTRO TO NLP

Name: Aditya Kumar Singh
Roll No.: 20217010 10

Q.1 Given probability distributions:

$A =$

	u_1	u_2	u_3
u_1	0.1	0.4	0.5
u_2	0.6	0.2	0.2
u_3	0.3	0.4	0.3

$b =$

	R	G	B
u_1	0.3	0.5	0.2
u_2	0.1	0.4	0.5
u_3	0.6	0.1	0.3

Initial probability $\pi_i = 0.33$

Compute the probabilities manually & report your observations at each 'λ' state.

A) Find the likelihood of the observation sequence: $\boxed{RRGG}$

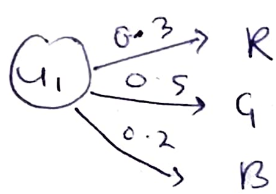
i.e., $P(RRGG \mid \lambda) = ?$

B) Find the Best State Sequence?

ANSWER:

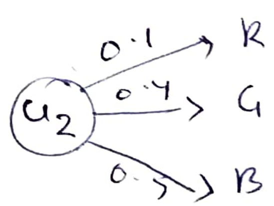
~~A) Using forward Method~~:

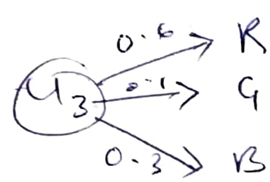
Here we have '3' hidden states: u_1, u_2, u_3 such that,



i.e., at state 'u_1' what is the emission probability of each letter R, G, B.

Similarly,





(1) → Now, using forward method, → we will calculate the "likelihood" of getting into one state & then what's the probability of obtaing 1st ~~step~~ term 'R' or 'G' or 'B' (i.e., the emission 2nd one) from that particular state.

(2) → Next, using the information of the current state (one of u_i's) & the output we'll estimated the likelihood of getting into one of u_1, u_2 & u_3 state & its corresponding output & so on.

(3) → Basically, (t+1)th state depends on (t)th state (i.e., only one step back) → hence this process from travereing from (t=1st) state to (t=4th) state, (in our case) is called induction process.

so let's begin:

i) Initialization step:

$\alpha_1(u_1)$ = At first ~~state~~ step → what's the ~~chance~~ probability of getting 'R' through 'u1' state.

$= \underbrace{\pi_{u_1}}_{\text{prob of getting into } u_1} \cdot b_{u_1}(\text{'R'}) = 0.33 \times 0.3$ (Being the first letter in RRGG)

$= \boxed{0.099}$

$\alpha_1(u_2)$ = " " through 'u2' state

$= \pi_{u_2} \cdot b_{u_2}(\text{'R'}) = 0.33 \times 0.1 = \boxed{0.033}$

$\alpha_1(u_3)$ = " " through 'u3' state

$= \pi_{u_3} \cdot b_{u_3}(\text{'R'}) = 0.33 \times 0.6 = \boxed{0.198}$

ii) INDUCTION step:

Once, we have ~~our~~ the information about which state ~~in~~ favour 'R' the most (or the desired outcome), we, using this, will decide which state (u_1, u_2, or u_3) will favour the next letter (i.e., 'R').

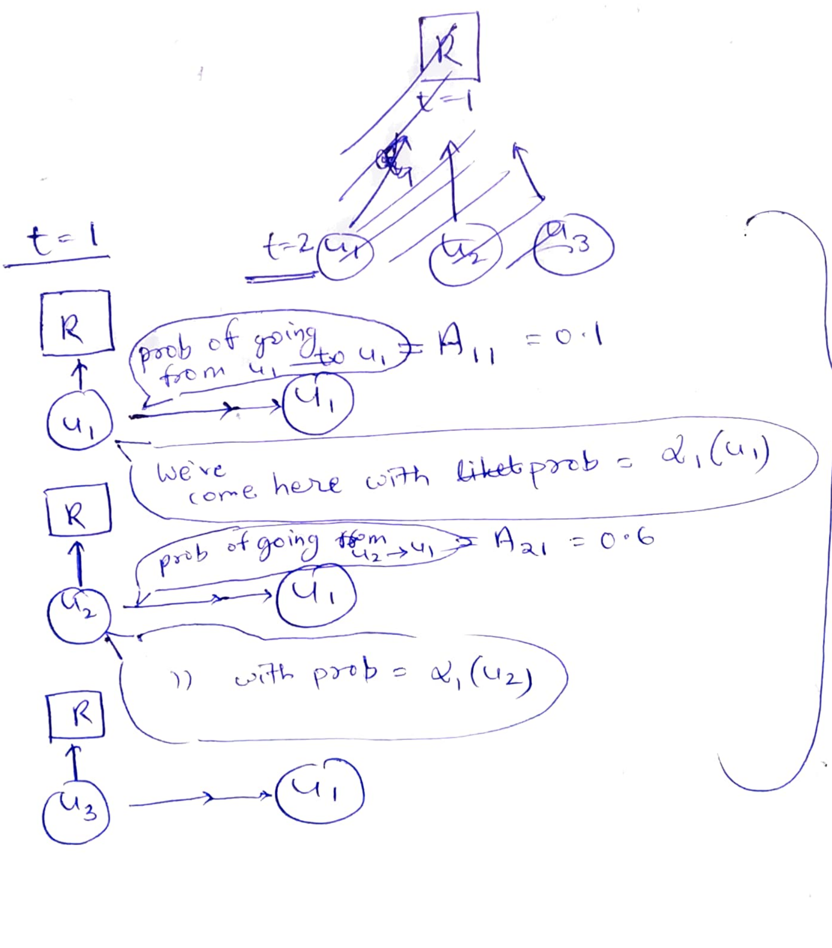
$$\alpha_{t+1}(u_j) = \left[\sum_{i=1}^{N} \alpha_t(u_i)\, \underbrace{A_{ij}}_{\text{state transition probability}}\right] b_{u_j}(\underbrace{\text{outcome}_{t+1}}_{\text{the letter at } (t+1)^{th} \text{ position}})$$

$\alpha_t(u_i)$: At step t' or time t', ~~what~~ which state has what likelihood for the desired outcome at that time step

plugging the values: (we get)

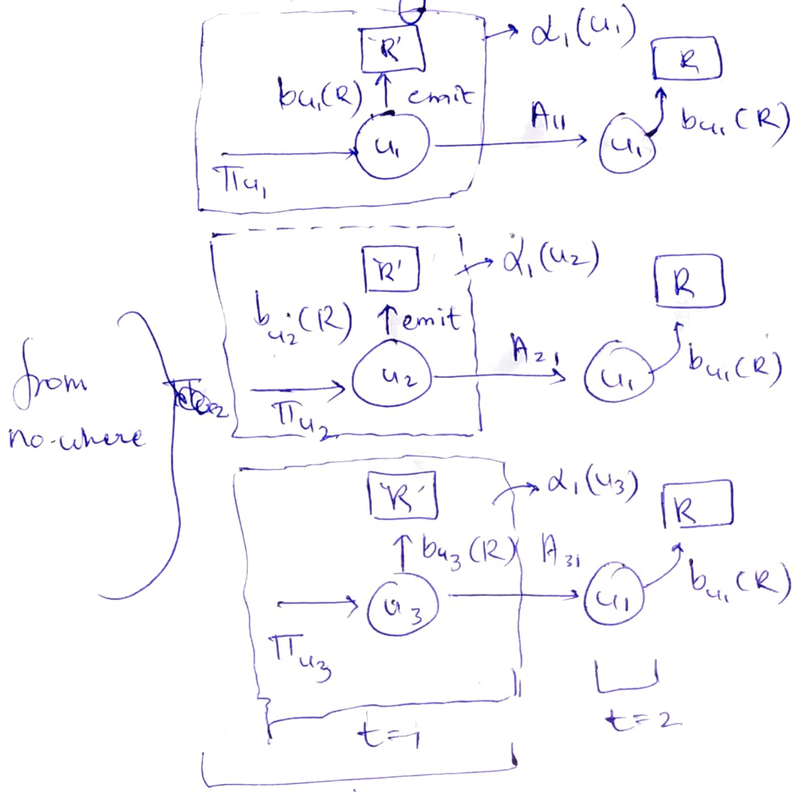
$$\alpha_2(u_1) = \underbrace{\left[\sum_{i=1}^{3} \alpha_1(u_i)\, A_{i1}\right]}_{\cancel{0.099}} b_{u_1}('R')$$

2nd ~~word~~ letter is 'R'



So we're combining all possible pathways possible to be at state u_1 in t=2.

• Next once we are at state "u_1", what is the chance of obtaining output 'R' there (i.e. what is $b_{u_1}('R')$)



t = 3 t = 4

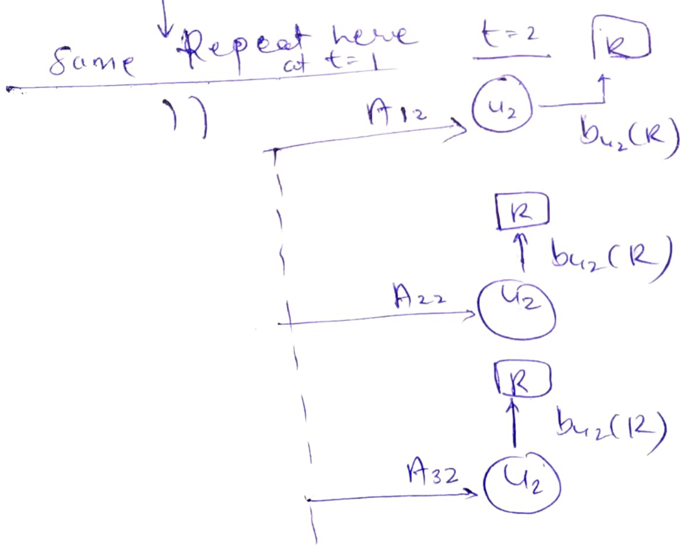
And this goes on

So basically, at each time step ~~...~~, t, there are '3' possible states $\boxed{u_1, u_2, \& u_3}$.

& total $3^4 = 81$ possible combinations possible:

u_1	u_1	u_1	u_7
t=1	t=2	t=3	t=4
u_1	u_1	u_1	u_2
⋮	⋮	⋮	⋮

same Repeat here at t=1))



t = 3 t = 4

~~And~~ And out of those, the ~~combination~~ ~~having total maximum total probability~~ path bearing maximum probability is chosen.

↳ And this is the whole idea.

So, back to computation:

$$\alpha_2(u_1) = \left[\alpha_1(u_1) A_{11} + \alpha_1(u_2) A_{21} + \alpha_1(u_3) A_{31}\right] b_{u_1}(R)$$

$$= \left[(0.099) \times (0.1) + 0.033 \times 0.6 + 0.198 \times 0.3\right] 0.3$$

$$= \boxed{0.02673}$$

$$\alpha_2(u_2) = \left[\alpha_1(u_1)A_{12} + \alpha_1(u_2)A_{22} + \alpha_1(u_3)A_{32}\right] b_{u_2}(R)$$

$$= \boxed{0.01254}$$

Similarly,

$$\alpha_2(u_3) = \left[\alpha_1(u_1)A_{13} + \alpha_1(u_2)A_{23} + \alpha_1(u_3)A_{33}\right] b_{u_3}(R)$$

$$= \boxed{0.0693}$$

$\alpha_3(u_1) = 0.0154935$

$\alpha_3(u_2) = 0.016368$

$\alpha_3(u_3) = 0.00366$

$\alpha_4(u_1) = \boxed{0.0062}$

$\alpha_4(u_2) = \boxed{0.0044}$

$\alpha_4(u_3) = \boxed{0.0012}$

iii) Termination step: (at $t=4$)

$$P(\text{output} \mid \lambda) = P(RRGG \mid \lambda)$$

$$= \sum_{i=1}^{3} \alpha_4(u_i)$$

$$= \alpha_4(u_1) + \alpha_4(u_2) + \alpha_4(u_3)$$

$$= \boxed{0.0118} \checkmark$$

B). To find the Best state sequence, we'll use "Viterbi" algorithm.

i) Initialization:

$$\underbrace{\delta_1(u_i)}_{\text{At timestep 1, what is the probability of getting 'R' at first step through } u_i \text{ state (= desired output)}} = \underbrace{\pi_{u_i}}_{\text{prob of entering into state } u_i} \; b_{u_i}(O_1) \quad 1 \leq i \leq N$$

$b_{u_i}(O_1)$ → probability of getting output as desired through u_i

In our case N = 3.

& state-variable $\varphi_1(u_i) = 0$

Computation :

$\delta_1(u_1) = \pi_{u_1} \, b_{u_1}(R) = 0.099$

$\delta_1(u_2) = \pi_{u_2} \, b_{u_2}(R) = 0.033$

$\delta_1(u_3) = \pi_{u_3} \, b_{u_3}(R) = 0.198$

$\varphi_1(u_1) = 0 = \varphi_1(u_2) = \varphi_1(u_3)$

ii) Recursively :

$$\delta_t(u_j) = \max_{1 \le i \le N} \left[\delta_{t-1}(u_i) A_{ij} \right] b_{u_j}(O_t)$$

where $2 \le t \le T$ (as $t-1$ is the initial step)

&

$$\varphi_t(u_j) = \operatorname*{argmax}_{1 \le i \le N} \left[\delta_{t-1}(u_i) A_{ij} \right]$$

basically this term dependent on $\boxed{\delta_{t-1}(u_i) A_{ij}}$ only, because $b_{u_j}(O_t)$ is common in all of them.

Computation :

$$\delta_2(u_1) = \max \left[\delta_1(u_1) A_{11}, \, \delta_1(u_2) A_{21}, \, \delta_1(u_3) A_{31} \right] b_{u_1}(R)$$

(2nd 'R' in RRGG)

$$= \max\left(\left[0.099 \times 0.1, \, 0.033 \times 0.6, \, 0.198 \times 0.3 \right] \cdot 0.3 \right)$$

$$= \max\left(\left[0.00297, \, 0.00594, \, 0.01782 \right] \right)$$

$$= 0.01782 \Rightarrow \varphi_2(u_1) = \operatorname{argmax}([\;]) = 3$$

$$\delta_2(u_2) = \max([\delta_1(u_1) A_{12}, \delta_1(u_2) A_{22}, \delta_1(u_3) A_{32}] \cdot b_{u_2}(R))$$

$$= 0.00792$$

$$\psi_2(u_2) = 3$$

Similarly:

$$\delta_2(u_3) = 0.03564, \quad \psi_2(u_3) = 3$$

t=3

$$\delta_3(u_1) = \max[\delta_2(u_1) A_{11}, \delta_2(u_2) A_{21}, \delta_2(u_3) A_{31}] \cdot b_{u_1}(G)$$

$$= \max[0.1782 \times 0.1 \times 0.5, \quad 0.00792 \times 0.6 \times 0.5, \quad 0.03564 \times 0.3 \times 0.5]$$

$$= \max[0.00891, 0.002376, 0.005346]$$

$$= 0.005346$$

$$\psi_3(u_1) = 3.$$

$$\delta_3(u_2) = 0.0057024, \quad \psi_3(u_2) = 3$$

$$\delta_3(u_3) = 0.0010692, \quad \psi_3(u_3) = 3$$

t=4

$$\delta_4(u_1) = 0.00171072, \quad \psi_4(u_1) = 2$$

$$\delta_4(u_2) = 0.00085536, \quad \psi_4(u_2) = 1$$

$$\delta_4(u_3) = 0.0002673, \quad \psi_4(u_3) = 1$$

iii) Termination:

$$p^* = \max_{1 \le i \le N} [\delta_T(u_i)]$$

T = last time step

$$= \max[\delta_4(u_1), \delta_4(u_2), \delta_4(u_3)] = 0.00171072$$

So Best path probability, out of all 81 combination for RRGG = $P^* = 0.00171072$

&

$q_4^* = \underset{1 \le i \le 3}{\operatorname{argmax}} [\delta_T(u_i)]$

$= \operatorname{argmax} [\delta_4(u_1), \delta_4(u_2), \delta_4(u_3)]$

$= \boxed{1}$.

Now, path sequence:

$q_t^* = \psi_{t+1}(q^*_{t+1})$

$\Rightarrow q_3^* = \psi_4(q_4^*) = \psi_4(1)$

or

$\psi_4(u_1) = \boxed{2}$

$\Rightarrow q_2^* = \psi_3(q_3^*) = \psi_3(2) = \psi_3(u_2)$

$= \boxed{3}$

$\Rightarrow q_1^* = \psi_2(q_2^*) = \psi_2(3) = \psi_2(u_3)$

$= 3$

∴ Best state sequence:

$q_1^* q_2^* q_3^* q_4^* = \boxed{3\ 3\ 2\ 1}$

x ——— o ——— x