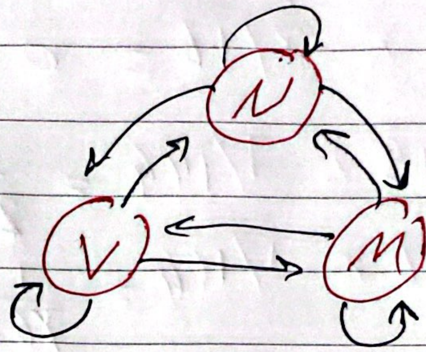
Sheet (5) Solutions
Sequence labeling

Q1: We have 3 POS tags which are → Noun (N), → Verb (V), → Modal (M)

Our corpus is as follows:

- Mary(N) Jane(N) Can(M) see(V) Will(N)
- Spot(N) will(M) see(V) Mary(N)
- Will(M) Jane(N) spot(V) Mary(N)?
- Mary(N) will(M) pat(V) Spot(N).



a) Calculate the emission probs: Markov State Graph

→ To compute the emission probability we use this formula → $p(w_i|t_i) = \frac{C(t_i, w_i)}{C(t_i)}$

بنشوف كام مره ال Tag ده جا مع ال word مقسوم على كام مرة ال Tag اتكرر.

Emission Table:- ✓

	Mary	Jane	Can	See	will	Spot	Pat
Noun	4/9	2/9	0	0	1/9	2/9	0
Verb	0	0	0	2/4	0	1/4	1/4
Model	0/4	0/4	1/4	0/4	3/4	0/4	0/4

Done ✓✓

b) Calculate the transition probabilities, adding the start tag & the end tag @ the start and end of each tag respectively.

→ To compute the transition probabilities we use the following formula → $P(t_i | t_{i-1}) = \frac{C(t_i, t_{i-1})}{C(t_i)}$

بنشوف كام مرة ال tag ti جا بعد ال tag ti-1 ونقسم على
العدد الكلي اللي ظهر فيه ال tag ti

Transition Matrix:-

	Noun	Verb	Model		</s>
<s>	3/4	0	1/4		
N	1/9	1/9	3/9	this is the π set	0/4
V	4/4	0	0		4/9
M	1/4	3/4	0/4		0
					0/4

متنساش تحط ال end tag في ال Table !

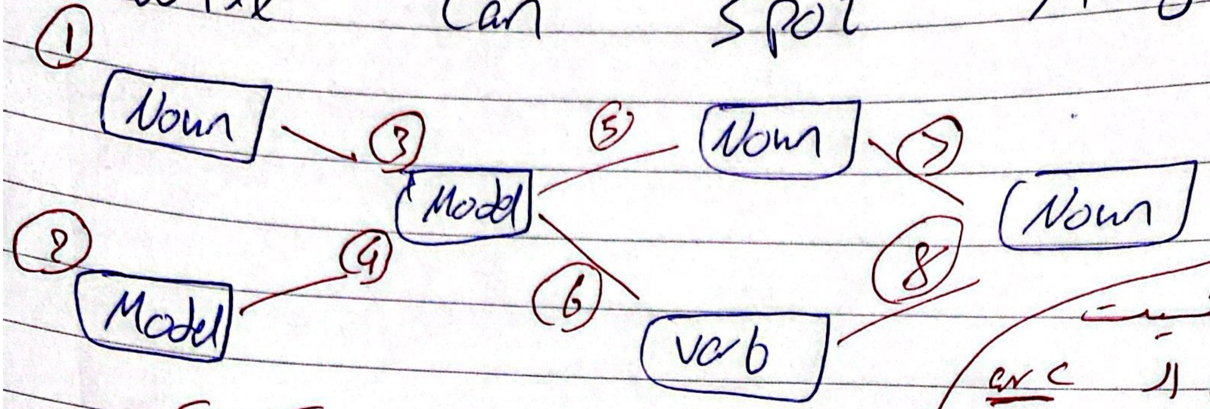
c) Using the Vetriby algorithm, mention the computed tag sequence for "will can spot Mary" → show the calculations.

Steps:

- Write the words as headings
- Under each word write the only possible states
- Do the computations!

Return the path.

Will Can Spot Mary



انا هنا نسيت
اضرب ف ال arc

Transition Matrix → Noun   Emission Matrix

① $P(\text{Noun} \mid \langle S \rangle) \times P(\text{will} \mid \text{Noun}) = 3/4 \times 1/9 = 1/12$

② $P(\text{Model} \mid \langle S \rangle) \times P(\text{will} \mid \text{Model}) = 1/4 \times 3/4 = 3/16$

③ $P(\text{Model} \mid \text{Noun}) \times P(\text{Can} \mid \text{Model}) = 3/9 \times 1/4 = 1/12 \times ①$

④ $P(\text{Model} \mid \text{Model}) \times P(\text{Can} \mid \text{Model}) = 0 \times 1/4 = 0 \times ②$

Now we will eleminate ④ & chose 3 only ✓✓

⑤ $P(\text{Noun} \mid \text{Model})\, P(\text{spot} \mid \text{Noun}) = 1/4 \times 2/9 = 1/18 \times ③$

⑥ $P(\text{Verb} \mid \text{Model})\, P(\text{spot} \mid \text{Verb}) = 3/4 \times 1/4 = 3/16 \times ③$

⑦ $P(\text{Noun} \mid \text{Noun})\, P(\text{Mary} \mid \text{Noun}) = 1/9 \times 4/9 = 4/81 \times ⑤$

⑧ $P(\text{Noun} \mid \text{verb})\, P(\text{Mary} \mid \text{Noun}) = 4/4 \times 4/9 = 4/9 \times ⑥$

Now we will elimenate 7 & chose 8

∴ Path is

Path evaluation :-

Because I did a mistake, lets revaluate the results

① & ② are the same values

③ = 1/12 x 1/12 = 1/144 ✓✓

④ = 0 x 3/16 = 0 ∴ will be eliminated

⑤ 1/18 x 1/144 = $\frac{1}{2592}$ = 3.858×10^{-4}

⑥ 3/16 x 1/144 = $\frac{1}{768}$ = 1.302×10^{-3}

⑦ $\frac{4}{81} \times \frac{1}{2592} = \frac{1}{52488} = 1.905 \times 10^{-5}$

⑧ $\frac{4}{9} \times \frac{1}{768} = \frac{1}{1728} = 5.787 \times 10^{-4}$

∴ 7 will be discarded

∴ the path will be

N	M	V.	N
will	Can	spot	Many

→ Correct & logical solution

Q2: a) Find the value of A

Usually we should multiply P(hen | PL) × P(PL | <s>)

→ But here we will apply addition instead

∴ A = 17 + 11 = 28

b) Find the value of B & C

→ Same logic ∴ B = P(vilar | PN) + min P(PN | all classes)

∴ P(PN | PL) = 3+28 (P(PN | PN) = 4+5)

P(PN | PP) = 4+22 P(PN | VB) = 2+23 ~~P(PN | EOS)~~

∴ B = 19 + 9 = 28 كل قيمة بجمعها مع ال cell اللي قبلها
اي اللي ناتج منها بجمع اصح.

∴ C = min P(</s> | all previous Pos) = 28

P(</s> | PL) + P(PL | PL) = 7 + 21 = 28 → ✓✓ = C

P(</s> | PN) + 35 = 43

P(</s> | PP) + 20 = 29

P(</s> | VB) + 36 = 43

Now the Tag sequence is → BOS X Y PL EOS

To evaluate Y, we know that we chose 21 for it then we need to apply the algorithm & select the best value.

$\therefore Y = P(\text{it} \mid PL) + \min \begin{cases} P(PL \mid PN) = 5 + 28 \\ P(PL \mid PP) = 12 + 27 \\ P(PL \mid PL) = 17 + 27 \\ P(PL \mid VB) = 3 + 14 = 17 \end{cases}$

$= 4 + 17 = 21$

$\therefore$ Y must be VB

Lets do the same for X

We know that we come from VB = 14

$\therefore X = P(\text{viler} \mid VB) + \min \begin{cases} P(VB \mid PL) + A = 5 + 28 = 33 \\ P(VB \mid PN) + 5 = 1 + 5 = 6 \\ P(VB \mid PP) + 2c = 0 \\ P(VB \mid VB) + 23 \end{cases}$

$\therefore X = 8 + 6 = 14$

$\therefore$ X must be PN

$\therefore$ the final sequence must be

BOS PN VB PL EOS #

✓

Start by highlighting the entities in the sentence and write O for any non entity and I for any entity

John	Alex	is	going	to	New	York	after	having	an	Appoint.	at
I	I	O	O	O	I	I	O	O	O	O	O

the	Artif.	Intel.	Cop.	in	Rome
O	I	I	I	O	I

IO Tagging

To generalize this into BIO, let the 1ST word of any entity be B

John	Alex	is	going	to	New	York	after	having	an	Appoint.	at
B	I	O	O	O	B	I	O	O	O	O	O

the	Artif.	Intel.	Cop.	in	Rome
O	B	I	I	O	B

BIO Tagging

To generalize this into BIOES, let the last word of any entity be E and any entity of one word only be S

John	Alex	is	going	to	New	York	after	having	an	Appoint.	at
B	E	O	O	O	B	E	O	O	O	O	O

the	Artif.	Intel.	Cop.	in	Rome
O	B	I	E	O	S

BIO Tagging