(2) a) P(w/sam)

P(I /Sam) = 3/5

P(</S> / Sam) = 2/5

Max(3/5, 2/5) = 3/5 (I)

∴ Next probable word is 'I'

I <S> Sam I

b) <S> Sam I do ....

P(w /do)

P(I /do) = 1/2

P(like/do) = 1/2

∴ Bothe the words are equally probable.

c) Sam I am Sam ...

P(w/sam)

P (I /sam) = 3/5

P (</S> / Sam) = 2/5

Max (3/5, 2/5) = 3/5 (I)

Hence, <S> Sam I am Sam I ...

d) <S> do I like ...

P(w/like)

P(Sam / like) = 1/3

P (</S>/like) = 2/3

Hence, probable word is "</S>"

(3) a) <S> Sam I do like </S>

P(a) = P(sam | <S>) * P(I | sam) * P (do| I) * P(like | do) * P(</S> | like)

$= \frac{3}{5} \times \frac{3}{5} \times \frac{1}{5} \times \frac{1}{2} \times \frac{2}{3}$

= 0.024

b) <s> Sam I am </s>

$P(b) = I(\text{Sam}|\langle s\rangle) * P(I|\text{Sam}) * P(\text{am}|I) * P(\langle/s\rangle|\text{am})$

$= \frac{3}{5} \times \frac{3}{5} \times \frac{2}{5} \times \frac{1}{2}$

$= 0.072$

c) <s> I do like Sam I am </s>

$P(c) = P(I|\langle s\rangle) * P(\text{do}|I) * P(\text{like}|\text{do}) * P(\text{Sam}|\text{like}) * P(I|\text{Sam}) * P(\text{am}|I) * P(\langle/s\rangle|\text{am})$

$= \frac{1}{5} \times \frac{1}{5} \times \frac{1}{2} \times 0 \times \frac{3}{5} \times \frac{2}{5} \times \frac{1}{2}$

$= 0$

∴ The sentence <s> Sam I am </s> is most probable.

Q.3] (a) Bi-gram LM with laplace smoothing

Words	Count
I	5
am	2
Sam	5
do	2
like	3

Total unique words = 6 (excluding <s>)

$P(\text{do}|I) = \frac{C(I|\text{do})+1}{C(I)+6} = \frac{1+1}{5+6} = \frac{2}{11}$

$P(\text{do}|\text{Sam}) = \frac{C(\text{Sam}|\text{do})+1}{C(\text{Sam})+6} = \frac{0+1}{5+6} = \frac{1}{11}$

$P(\text{Sam}|\langle s\rangle) = \frac{C(\langle s\rangle|\text{Sam})+1}{C(\langle s\rangle)+6} = \frac{3+1}{5+6} = \frac{4}{11}$

$$P(I|sam) = \frac{C(sam|I)+1}{C(sam)+6} = \frac{3+1}{5+6} = \frac{4}{11}$$

$$P(I|do) = \frac{C(do|I)+1}{C(do)+6} = \frac{1+1}{2+6} = \frac{1}{4}$$

$$P(like|I) = \frac{C(I|like)+1}{C(I)+6} = \frac{2+1}{5+6} = \frac{3}{11}$$

② a) $P = P(do|<s>) \cdot P(sam|do) \cdot P(I|sam) \cdot P(like|I)$

$$= \frac{1+1}{5+6} \cdot \frac{0+1}{2+6} \cdot \frac{3+1}{5+6} \cdot \frac{2+1}{5+6}$$

$$= 2.25 \times 10^{-3}$$

b) $P = P(sam|<s>) \cdot P(do|sam) \cdot P(I|do) \cdot P(like|I)$

$$= \frac{3+1}{5+6} \cdot \frac{0+1}{5+6} \cdot \frac{1+1}{2+6} \cdot \frac{2+1}{5+6}$$

$$= 2.25 \times 10^{-3}$$

∴ The two sequences are equally probable

Q.4] ~~Case 1 (H) C(R) = C(I) C(0) = 1~~

	Null	H	U	N	D	A	I
Null	(0)	1	2	3	4	5	6
H	1	(0)	1	2	3	4	5
O	2	1	(1)	1	2	3	4
N	3	2	2	(1)	2	3	4
D	4	3	3	2	(1)	2	3
A	5	4	4	3	2	(1)	(2)
I							

HUNDAI Replace (U, O)
HONDAI Delete (I)

~~Minimum Edit Distance is 2~~

Case 1: $C(R) = C(I) = C(D) = 1$

∴ Minimum Edit distance $= 1 * C(R) + 1 * C(O)$
$= 1 \times 1 + 1 \times 1$
$= 2$

Case 2: $C(I) = C(O) = 1$ & $C(R) = 2$

∴ Minimum Edit Distance $= 1 * C(R) + 1 * C(D)$
$= 1 \times 2 + 1 \times 1$
$= 3$

Q.5] a) Bigram, $P(t_i, t_{i-1}) = \frac{C(t_{i-1}, t_i)}{C(t_{i-1})}$

Trigram, $P(t_i, t_{i-1}, t_{i-2}) = \frac{C(t_{i-2}, t_{i-1}, t_i)}{C(t_{i-2}, t_{i-1})}$

(i) $P(JJ|DT) = \frac{4}{12} = 0.333$

(ii) ~~$P(NN|DT, JJ)$~~
$P(VB|TO) = \frac{2}{2} = 1$

(iii) $P(NN|DT, JJ) = \frac{4}{4} = 1$

b) $P(w_i|t_i) = \frac{C(t_i, w_i)}{C(t_i)}$

(i) $P(go|VB) = \frac{2}{6} = 0.333$

(ii) $P(fish|NN) = \frac{7}{15} = 0.466$

Q.6] $P(\text{the light book}|\text{Det J J NN})$

$= P(the|DT) * P(light|JJ) \cdot P(book|NN) \cdot P(DT|start) \cdot$

$P(JJ|DT) \cdot P(NN|JJ)$

$= 0.3 \times 0.002 \times 0.003 \times 0.45 \times 0.3 \times 0.2$

$= 4.86 \times 10^{-8}$