B]

Q.1] a] $q_0 = \varepsilon \longrightarrow$ (1)

$q_1 = q_0 a + q_2 a \longrightarrow$ (2)

$q_2 = q_1 b + q_3 b \longrightarrow$ (3)

$q_3 = q_1 a \longrightarrow$ (4)

(2) can be re-written as -

$q_1 = a + q_2 a \longrightarrow$ (5)

Now (3) $\Rightarrow$

$q_2 = q_1 b + q_3 b$

from (4) $\rightarrow$

$q_2 = q_1 b + (q_1 a) b$

from (5) $\rightarrow$

$q_2 = (a + q_2 a) b + ((a + q_2 a) a) b$

$q_2 = ab + q_2 ab + aab + q_2 aab$

$q_2 = (ab + aab) + q_2 (ab + aab)$

by Arden's theorem,

if $R = Q + PR$, then, $R = QP*$

$\therefore q_2 = (ab + aab)(ab + aab)*$

$$q_2 = (aa?b) \cdot (aa?b)*$$
$$q_2 = (aa?b)+$$

$$\therefore \boxed{q_2 = (aa?b)+}$$

Q.1 b] $q_0 = \varepsilon \longrightarrow ①$
$q_1 = q_0 a + q_2 a \longrightarrow ②$
$q_2 = q_1 b + q_3 a \longrightarrow ③$
$q_3 = q_1 b \longrightarrow ④$

③ $\Rightarrow$

$q_2 = q_1 b + (q_1 b) a$ — from ④
$q_2 = (a + q_2 a) b + ((a + q_2 a) b) a$ — from ① & ②

$q_2 = ab + q_2 ab + aba + q_2 aba$
$q_2 = (ab + aba) + q_2 (ab + aba)$
by Arden's thm., -
$q_2 = (ab + aba) \cdot (ab + aba)*$
$q_2 = (ab + aba)+$
$q_2 = (aba?)+$

$$\therefore \boxed{q_2 = (aba?)+}$$

Q.2 Bigram Probabilities

"The Fed chairman warned that the board's decision is bad"

Prob (given sentence) =

P(Fed | The) × P(chairman | Fed) × P(warned | chairman) × P(that | warned) × P(the | that) × P(board | the) × P('s | board) × P(decision | 's) × P(is | decision) × P(bad | is)

i] No smoothing ⇒

P(given sentence) = 0.01315 × 0.2105 × 0 × 0.33 × 0.2023 × 0.1006 × 0.0464 × 0.025 × 0.0588 × 0

$\boxed{P(\text{given sentence})_{\text{No smoothing}} = 0}$ //

ii] Add One Smoothing ⇒

Let's calculate the unobserved bigrams first :

$$P(\text{warned} \mid \text{chairman}) = \frac{0+1}{C(\text{chairman}) + V}$$

$$= \frac{1}{554 + 5606}$$

$$\therefore P(\text{warned} \mid \text{chairman}) = 1.6233 \times 10^{-4}$$

$$P(\text{bad} \mid \text{is}) = \frac{0+1}{C(\text{is}) + V}$$

$$= \frac{1}{187 + 5606}$$

$$\therefore P(\text{bad} \mid \text{is}) = 1.7262 \times 10^{-4}$$

$$P_{\text{Add One}}(\text{given sentence}) = 5.2101 \times 10^{-4} \times 8.88 \times 10^{-4} \times 1.623 \times 10^{-4} \times \cancel{0.00406}\ 3.565 \times 10^{-4}$$

$$\times 0.00903 \times 0.0206 \times 0.00301$$

$$\times 0.00183 \times 3.556 \times 10^{-4} \times 1.726 \times 10^{-4}$$

$$= 5.028 \times 10^{-30}$$

$$\boxed{P_{\text{Add One Smoothing}}(\text{given sentence}) = \cancel{5.028 \times 10^{-30}}\ 1.694 \times 10^{-30}}$$

iii] Good Turing Discount :

$$P(\text{given sentence}) = 3.65 \times 10^{-5} \times 9.106 \times 10^{-5}$$

$$\times 0 \times 7.62 \times 10^{-6} \times 0.0019 \times 0$$

$$\times 3.251 \times 10^{-4} \times 2.529 \times 10^{-4} \times$$

$$7.622 \times 10^{-6} \times 0$$

$$= 0.00$$

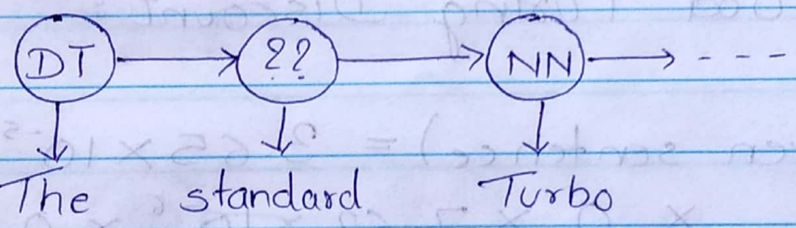
$$\boxed{P_{\text{Good Turing}}(\text{given sentence}) = 0}$$

Q.3b

We have two tags (??) to compute in the sentence.

According to the bigram model, the only probabilities being affected would be the neighbouring ones.
We therefore, will consider only the neighbours of ??

ai] First, let's consider "standard"



We have two tags for standard in our a (word | tag) bigrams - JJ & NN.
Let's calculate all the required probabilities:

$P(\text{standard} \mid JJ) = 7.745 \times 10^{-4}$

$P(\text{standard} \mid NN) = 6.3051 \times 10^{-4}$

$P(JJ \mid DT) = 0.22808$

$P(NN \mid DT) = 0.50946$

$P(NN \mid JJ) = 0.54350$

$P(NN \mid NN) = 0.17123$

Q.3.b]

$\therefore P(\text{standard} \mid JJ) \times P(JJ \mid DT) \times P(NN \mid JJ)$

$= \underline{9.6008 \times 10^{-5}}$

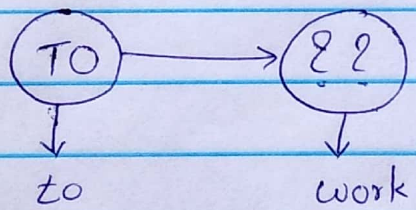
&

$P(\text{standard} \mid NN) \times P(NN \mid DT) \times P(NN \mid NN)$

$= \underline{5.4985 \times 10^{-5}}$

$\therefore$ We choose <u>JJ</u> as the tag for <u>standard</u>. //

ii] Similarly for "work"



We have 3 tags for work - NN, VB & VBP

$P(\text{work} \mid NN) = 0.002251$

$P(\text{work} \mid VB) = 0.00350$

$P(\text{work} \mid VBP) = 0.0114$

$P(NN \mid TO) = 0.03288$

$P(VB \mid TO) = 0.6357$

$P(VBP \mid TO) = 0.0$

$\therefore P(work|NN) \times P(NN|TO) = 7.4012 \times 10^{-5}$

$P(work|VB) \times P(VB|TO) = 2.224 \times 10^{-3}$

$P(work|VBP) \times P(VBP|TO) = 0$

$\therefore$ We choose VB as the tag for work.