1) $S = \{1, 2, 3\}$ → state space

$\mathcal{Y} = \{0, 1\}$ → output alphabet

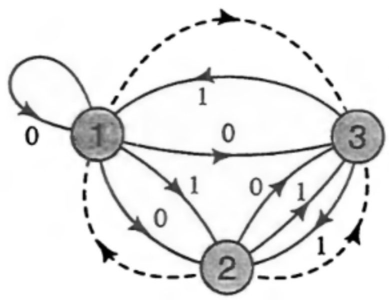


Figure 2.8
Addition of null transitions to the HMM of figure 2.3

Transition Probabilities :-

1) output producing arcs
2) Null arcs

$\Big\}$ both sum upto 1

output producing ↑ $\qquad\qquad$ Null ↑

$$p(s'|s) = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & \frac{1}{3} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{bmatrix} \quad \text{and} \quad q(s'|s) = \begin{bmatrix} 0 & 0 & \frac{1}{6} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}, \quad s, s' \in \mathcal{S}.$$

Emission probabilities :-

$q(1 \mid s \to s') = 1 - q(0 \mid s \to s')$

→ one matrix its 0 other is 1

→ " " " ½ " " ½

$$q(\mathbf{0}|s \to s') = \begin{bmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad q(\mathbf{1}|s \to s') = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 1 & 1 & \frac{2}{3} \\ 1 & 1 & 1 \end{bmatrix},$$

Initial state :-

$S_0 = 1$

(a) Draw state Diagram

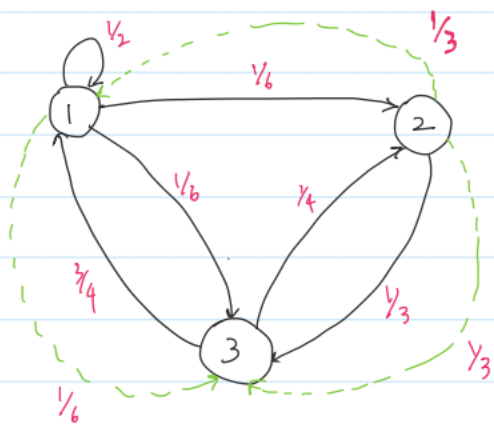
attach state labels

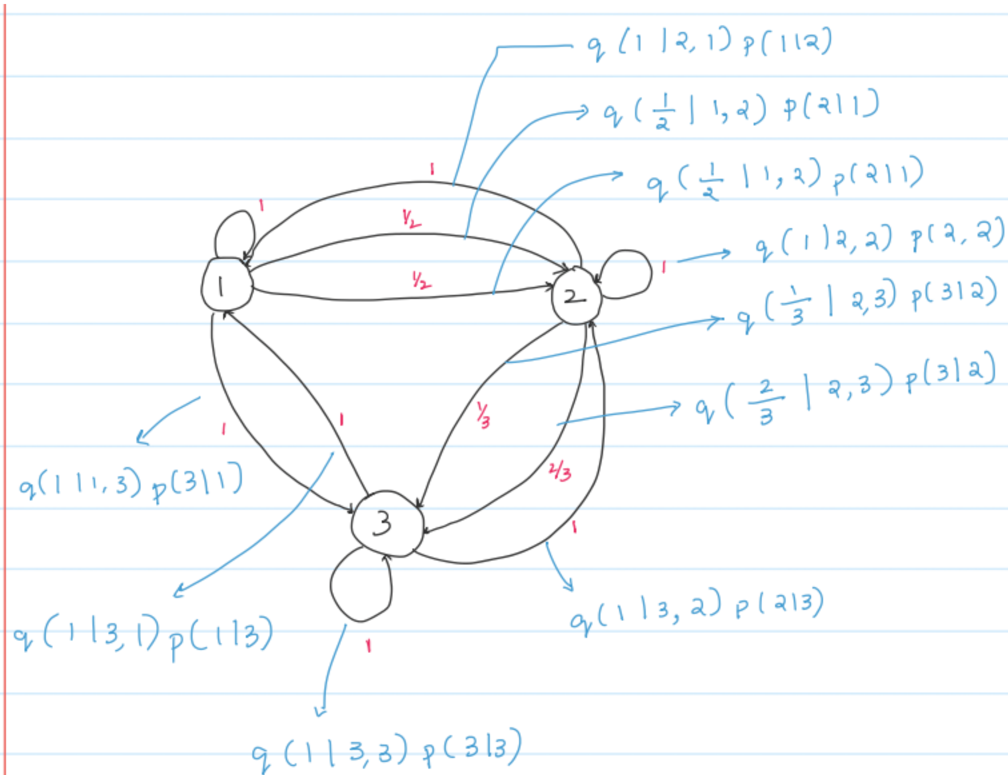
attach probabities $r(y, s' | s)$ or $q(s'|s)$ to all arcs

probability of non-null transition

probabity of null transition

attach output labels to all null arcs.



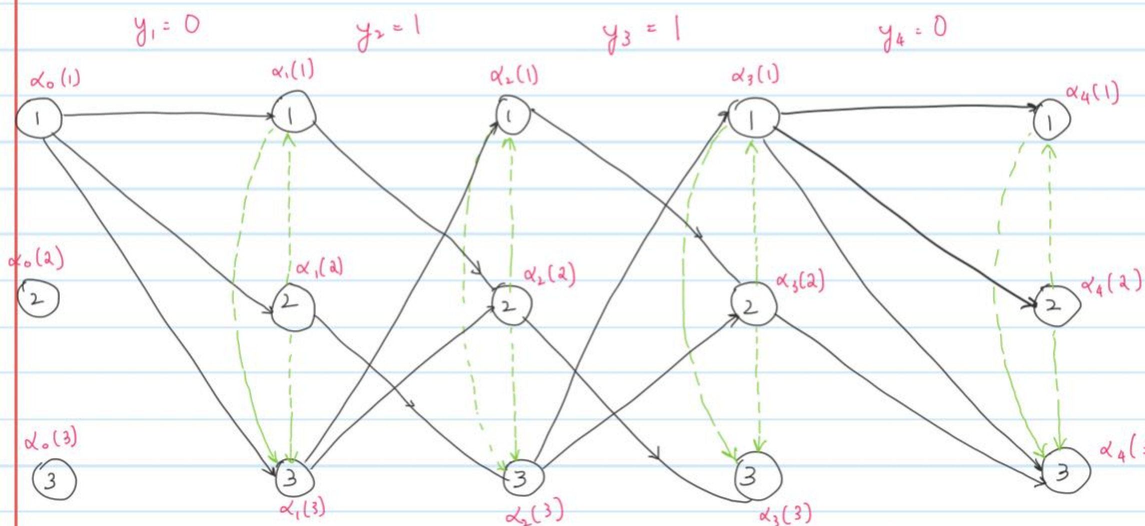


b) trellis $y_1 y_2 y_3 y_4 = 0110$

only output produing

b) trellis $y_1 y_2 y_3 y_4 = 0110$

only output producing



c) Compute forward probabity $\alpha_i(s)$?

given

$$A = \begin{bmatrix} p(1|1) & p(2|1) & p(3|1) \\ p(1|2) & p(2|2) & p(3|2) \\ p(1|3) & p(2|3) & p(3|3) \end{bmatrix} \quad p(s'|s) = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & \frac{1}{3} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} q(1|1) & q(2|1) & q(3|1) \\ q(1|2) & q(2|2) & q(3|2) \\ q(1|3) & q(2|3) & q(3|3) \end{bmatrix} \quad q(s'|s) = \begin{bmatrix} 0 & 0 & \frac{1}{6} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

$$B(0) = \begin{bmatrix} q(Y=0|1,1) & q(Y=0|2,1) & q(Y=0|3,1) \\ q(Y=0|1,2) & q(Y=0|2,2) & q(Y=0|3,2) \\ q(Y=0|1,3) & q(Y=0|2,3) & q(Y=0|3,3) \end{bmatrix} \quad q(\mathbf{0}|s \to s') = \begin{bmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

$$B(1) = \begin{bmatrix} q(Y=1|1,1) & q(Y=1|2,1) & q(Y=1|3,1) \\ q(Y=1|1,2) & q(Y=1|2,2) & q(Y=1|3,2) \\ q(Y=1|1,3) & q(Y=1|2,3) & q(Y=1|3,3) \end{bmatrix} \quad q(\mathbf{1}|s \to s') = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 1 & 1 & \frac{2}{3} \\ 1 & 1 & 1 \end{bmatrix}$$

$p(s_0 = 1) = 1$

$$\alpha_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$p(Y_1 = 0, S_1 = 1 \mid S_0)$
$= \alpha_0(1)\ A_{1,1}\ B_{1,1}(0)$
$= (1)(1/2)(1)$
$= 1/2$

$p(Y_1 = 0, S_1 = 2 \mid S_0)$
$= \alpha_0(1)\ A_{1,2}\ B_{1,2}(0)$
$= (1)(1/6)(1/2)$
$= 1/12$

$p(Y_1 = 0, S_1 = 3 \mid S_0)$
$= \alpha_0(1)\ A_{1,3}\ B_{1,3}(0)$
$= (1)(1/6)(1)$
$= 1/6$

$$\alpha_1' = \begin{pmatrix} 1/2 \\ 1/12 \\ 1/6 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.0833 \\ 0.16667 \end{pmatrix} \rightarrow \text{without null}$$

$\alpha_1(2) = \alpha_1'(2) = 1/12$

$\alpha_1(1) = \alpha_1'(1) + \alpha_1(2)\, V_{2,1}$
$= 1/2 + 1/12\,(1/3)$
$= 1/2 + 1/36$
$= 0.5278$

$\alpha_1(3) = \alpha_1'(3) + \alpha_1(2)\, V_{2,3} + \alpha_1(1)\, V_{1,3}$
$= 1/6 + (1/12)(1/3) + 0.5278\,(1/6)$
$= \frac{1}{6} + \frac{1}{36} + 0.0879$
$= 0.2824$

$$\alpha_1 = \begin{pmatrix} 0.527 \\ 0.083 \\ 0.2824 \end{pmatrix}$$

$$\alpha_2' = \begin{pmatrix} \alpha_1(3)\, A_{3,1}\ B_{3,1}(1) \\ \alpha_1(1)\, A_{1,2}\ B_{1,2}(1) + \alpha_1(3)\, A_{3,2}\ B_{3,2}(1) \\ \alpha_1(2)\, A_{2,3}\ B_{2,3}(1) \end{pmatrix}$$

$$\alpha_2' = \begin{pmatrix} 0.2824\,(3/4)(1) \\ 0.527\,(1/6)(1/2) + 0.2824\,(1/4)(1) \\ 0.083\,(1/3)(2/3) \end{pmatrix} = \begin{pmatrix} 0.2118 \\ 0.114 \\ 0.0185 \end{pmatrix}$$

$$\alpha_2(2) = \alpha_2'(2) = 0.114$$

$$\alpha_2(1) = \alpha_2'(1) + \alpha_2(2)\, V_{2,1}$$
$$= 0.207 + 0.114\,(1/3)$$
$$= 0.25$$

$$\alpha_2(3) = \alpha_2'(3) + \alpha_2(2)\, V_{2,3} + \alpha_2(1)\, V_{1,3}$$
$$= 0.0185 + 0.114\,(1/3) + 0.25\,(1/6)$$
$$= 0.09817$$

$$\alpha_2 = \begin{pmatrix} 0.25 \\ 0.114 \\ 0.09817 \end{pmatrix}$$

$$\alpha_3' = \begin{pmatrix} \alpha_2(3)\, A_{3,1}\ \ B_{3,1}(1) \\ \alpha_2(1)\, A_{1,2}\ \ B_{1,2}(1) + \alpha_2(3)\, A_{3,2}\ B_{3,2}(1) \\ \alpha_2(2)\ A_{2,3}\ \ B_{2,3}(1) \end{pmatrix}$$

$$= \begin{pmatrix} 0.098\ (3/4)(1) \\ 0.25\ (1/6)(1/2) + 0.09817 (1/4)(1) \\ 0.114\ (1/3)(2/3) \end{pmatrix}$$

$$= \begin{pmatrix} 0.07 \\ 0.045 \\ 0.0253 \end{pmatrix}$$

$$\alpha_3(2) = \alpha_3'(2) = 0.045$$

$$\alpha_3(1) = \alpha_3'(1) + \alpha_3(2)\, V_{2,1}$$

$$= 0.07 + 0.045\,(1/3)$$

$$= 0.085$$

$$\alpha_3(3) = \alpha_3'(3) + \alpha_3(2)\,V_{2,3} + \alpha_3(1)\,V_{1,3}$$

$$= 0.0253 + 0.085\,(1/3) + 0.045\,(1/6)$$

$$= 0.061$$

$$\alpha_3 = \begin{pmatrix} 0.085 \\ 0.045 \\ 0.061 \end{pmatrix}$$

$$\alpha_4' = \begin{pmatrix} \alpha_3(1)\, A_{1,1}\, B_{1,1}(0) & \\ \alpha_3(1)\, A_{1,2}\, B_{1,2}(0) & + \; \alpha_3(1)\, A_{1,3}\, B_{1,3}(0) \\ \alpha_3(2)\, A_{2,3}\, B_{2,3}(0) & \end{pmatrix}$$

$$= \begin{pmatrix} 0.085\,(1/2)\,(1) & \\ 0.085\,(1/6)\,(1/2) & + \; 0.085\,(1/6)(1) \\ 0.045\,(1/3)\,(1/3) & \end{pmatrix}$$

$$= \begin{pmatrix} 0.0425 \\ 0.02125 \\ 0.005 \end{pmatrix}$$

$$\alpha_4(2) = \alpha_4'(2) = 0.02125$$

$$\alpha_4(1) = \alpha_4'(1) + \alpha_4(2)\, V_{2,1}$$

$$= 0.0425 + 0.021\,(1/3)$$

$$= 0.0495$$

$$\alpha_4(3) = \alpha_4'(3) + \alpha_4(2)\, V_{2,3} + \alpha_4(1)\, V_{1,3}$$

$$= 0.005 + 0.021\,(1/3) + 0.0495\,(1/6)$$

$$= 0.0203$$

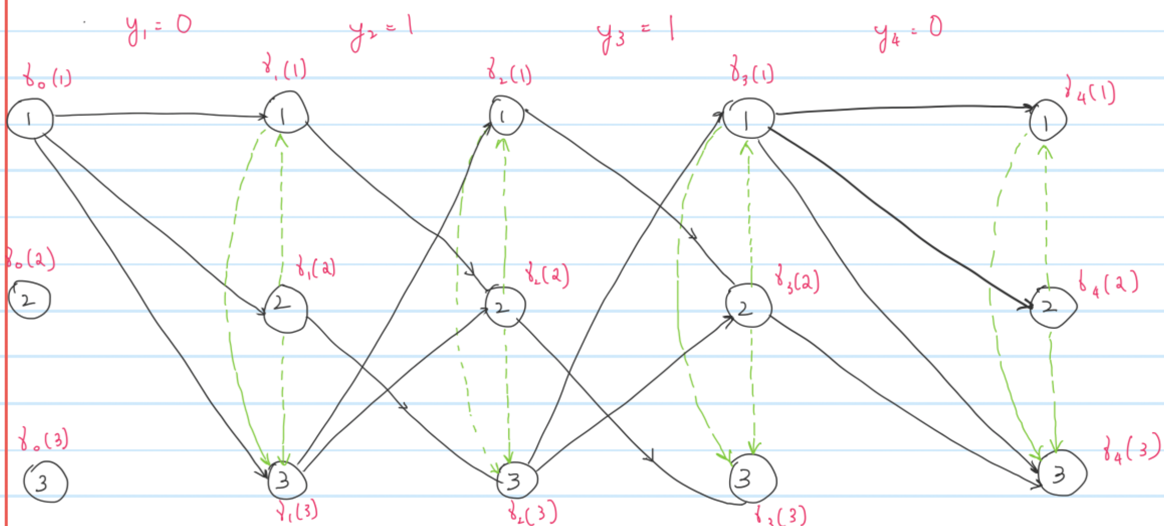
$$\alpha_4 = \begin{pmatrix} 0.0495 \\ 0.02125 \\ 0.0203 \end{pmatrix}$$

d) Marginal probability

$$P(y_1 y_2 y_3 y_4 = 0110 \mid s_0 = 1) = \sum_{i=1}^{3} \alpha_4(i)$$

$$= 0.0495 + 0.02125$$
$$+ 0.0203$$
$$= 0.110105$$

e) Viterbi probability $\delta_i(s)$



$\gamma_0(1) = 1$

$$\gamma_1 = \begin{pmatrix} A_{11} \cdot B_{11}(0) \ , \ A_{12} \cdot B_{12}(0) \cdot V_{21} B_{21}(0) \\ A_{12} \cdot B_{12}(0) \\ A_{13} \cdot B_{13}(0) \ , \ A_{12} B_{12}(0) \cdot V_{23} B_{23}(0) \ , \ A_{11} B_{11}(0) \cdot V_{13} B_{13}(0) \end{pmatrix}$$

$$= \max \begin{pmatrix} 1/2 \ , \ 0 \\ 1/6 \cdot 1/2 \\ 1/6 \ , \ \frac{1}{12} \cdot \frac{1}{3} \cdot \frac{1}{3} \ , \ \frac{1}{2} \frac{1}{6} \ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/6 \times \frac{1}{2} \\ 1/6 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.0833 \\ 0.16667 \end{pmatrix}$$

$$\gamma_2(1) = \max \left(A_{13} \cdot B_{13}(0)\ A_{31}\ B_{31}(1) \ , \ \gamma_1(3) \cdot A_{32}\ B_{32}(1) \cdot V_{21}\ B_{21}(1) \right)$$

$$\max \left(\frac{1}{6} \times 1 \times \frac{3}{4} \times 1 \ , \ \frac{1}{6} \cdot \frac{1}{4} \cdot 1 \cdot 0 \right)$$

$$\max \left(\frac{1}{8} \ , 0 \right)$$

$$= 1/8 = 0.125$$

$$\gamma_2(2) = \max \left(\gamma_1(1) \cdot A_{12} \cdot B_{12}(1) + \gamma_1(3) \cdot A_{32} \cdot B_{32}(1) \right)$$

$$= \max \left(\frac{1}{2} \times \frac{1}{6} \times \frac{1}{2} \ , \ \frac{1}{6} \times \frac{1}{4} \times 1 \right)$$

$$= \max \left(\frac{1}{24} \ , \ \frac{1}{24} \right)$$

$$= 0.04166$$

$$\gamma_2(3) = \max\left(\gamma_1(2) \cdot A_{2,3}\, B_{2,3}(1),\ \gamma_1(3)\, A_{32}\, B_{32}(1)\, V_{23},\ \gamma_1(3)\, A_{31}\, B_{31}(1)\, V_{13}\right)$$

$$\max\left(\frac{1}{\cancel{12}_{6}} \times \frac{1}{3} \times \frac{\cancel{2}}{3},\ \frac{1}{\cancel{6}_{3}} \times \frac{1}{4}\ 1 \times \frac{1}{3} \times \frac{\cancel{2}}{3},\ \frac{1}{6} \times \frac{\cancel{3}}{\cancel{4}_{2}} \times 1 \times \frac{1}{6}\right)$$

$$\max\left(\frac{1}{54},\ \frac{1}{108},\ 0\right)$$

$$= (0.0185,\ 0.02083,\ 0)$$

$$\gamma_2 = \begin{pmatrix} 0.125 \\ 0.04166 \\ 0.02083 \end{pmatrix}$$

$$\gamma_3(1) = \max\left(\gamma_2(3) \cdot A_{3,1}\, B_{3,1}(1),\ \gamma_2(3) \cdot A_{3,2}\, B_{3,2}(1),\ V_{2,1}\right)$$

$$= \max\left(\frac{1}{\cancel{54}_{18}} \times \frac{\cancel{3}}{4} \times 1,\ \frac{1}{54} \cdot \frac{1}{4} \cdot 1,\ 0\right)$$

$$= \frac{1}{72} = 0.0138$$

$$\gamma_3(2) = \max\left(\gamma_2(1) \cdot A_{12}\, B_{12}(1),\ \gamma_2(3) \cdot A_{32}\, B_{32}(1)\right)$$

$$= \max\left(\frac{1}{8} \cdot \frac{1}{6} \cdot \frac{1}{2},\ \frac{1}{54} \cdot \frac{1}{4} \cdot 1\right)$$

$$= \max\left(\frac{1}{96},\ \frac{1}{216}\right)$$

$$= \frac{1}{96} = 0.0104$$

$$\gamma_3(3) = \gamma_2(2) \cdot A_{2,3}\, B_{2,3}(1)$$

$$= \frac{1}{\cancel{24}_{12}} \cdot \frac{1}{3} \cdot \frac{\cancel{2}}{3} = \frac{1}{108} = 0.009$$

$$\gamma_3 = \begin{pmatrix} 1/72 \\ 1/96 \\ 1/108 \end{pmatrix} = \begin{pmatrix} 0.0138 \\ 0.0104 \\ 0.00925 \end{pmatrix}$$

$$\gamma_4(1) = \gamma_3(1) \cdot A_{11}\, B_{11}(0)$$

$$= 0.0138 \cdot 0.5 \cdot 1 = 0.0078$$

$$\gamma_4(2) = \gamma_3(1) \cdot A_{12}\, B_{12}(0)$$

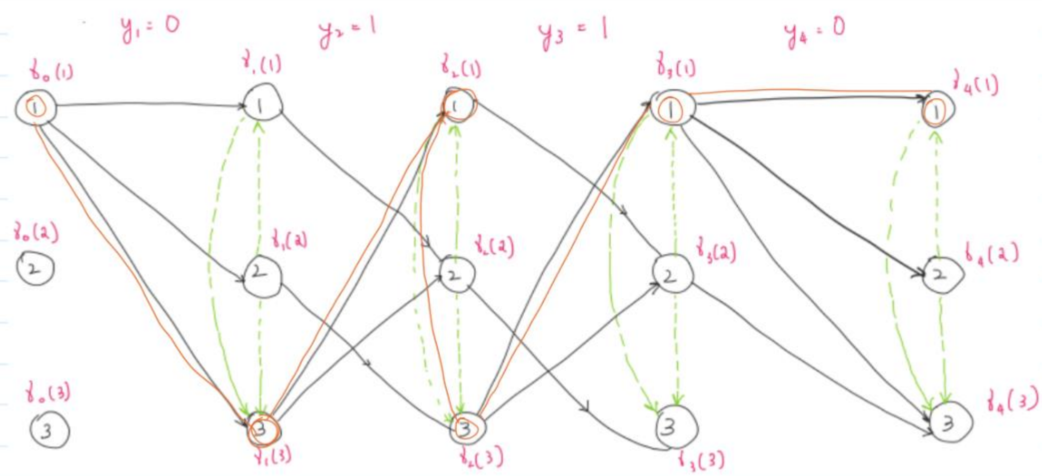
$$= 0.0138 \cdot \frac{1}{12} = 0.0015$$

$$\gamma_4(3) = \max\left(\gamma_3(2) \cdot A_{23}\, B_{23}(0) + \gamma_3(1)\, A_{13}\, B_{13}(0)\right)$$

$$= 0.002$$

$$\gamma_4 = \begin{pmatrix} 1/128 \\ 1/768 \\ 1/384 \end{pmatrix} = \begin{pmatrix} 0.0078125 \\ 0.00130208 \\ 0.002641 \end{pmatrix}$$

f) likelist path



2) proof :-

Certainly! Let's delve into the properties of Hidden Markov Models (HMMs) and the forward probabilities ($\alpha_i(s)$) to show how we can efficiently compute the marginal probability of an observation sequence.

The marginal probability of an observation sequence O in an HMM is given by:

$P(O) = \sum_s P(O, q_T = s)$

where q_T represents the state at time T and s ranges over all possible states. Now, using the definition of conditional probability, we can express $P(O, q_T = s)$ as the product of the probability of the observation sequence given the state ($P(O|q_T = s)$) and the probability of being in state s at time T ($P(q_T = s)$):

$P(O, q_T = s) = P(O|q_T = s) \cdot P(q_T = s)$

Now, we can express $P(O|q_T = s)$ in terms of the observed sequence $O_{1:T}$ and the state at time T:

$P(O|q_T = s) = P(o_T|q_T = s) \cdot P(O_{1:T-1}|q_T = s)$

Here, $P(o_T|q_T = s)$ is the emission probability of observing o_T given the state s, and $P(O_{1:T-1}|q_T = s)$ is the probability of the partial observation sequence $O_{1:T-1}$ given the state s at time T.

Now, substitute this back into the expression for $P(O, q_T = s)$:

$P(O, q_T = s) = P(o_T|q_T = s) \cdot P(O_{1:T-1}|q_T = s) \cdot P(q_T = s)$

Now, the marginal probability $P(O)$ can be obtained by summing over all possible states:

$P(O) = \sum_s P(O, q_T = s)$

$P(O) = \sum_s P(o_T|q_T = s) \cdot P(O_{1:T-1}|q_T = s) \cdot P(q_T = s)$

This expression involves the forward probabilities ($\alpha_i(s)$) which are defined as the probability of the partial observation sequence $O_{1:i}$ and being in state s at time i. Specifically, $\alpha_i(s) = P(O_{1:i}, q_i = s)$.

Now, we can express $P(O_{1:T-1}|q_T = s)$ in terms of the forward probabilities at time $T-1$ using the recursive relationship:

$P(O_{1:T-1}|q_T = s) = \sum_{s'} P(q_{T-1} = s'|q_T = s) \cdot P(O_{1:T-1}|q_{T-1} = s')$

Substitute this back into the expression for $P(O)$ and continue the recursion until you reach $\alpha_1(s)$. This recursive computation allows for an efficient computation of the marginal probability of the observation sequence using the forward probabilities in an HMM.