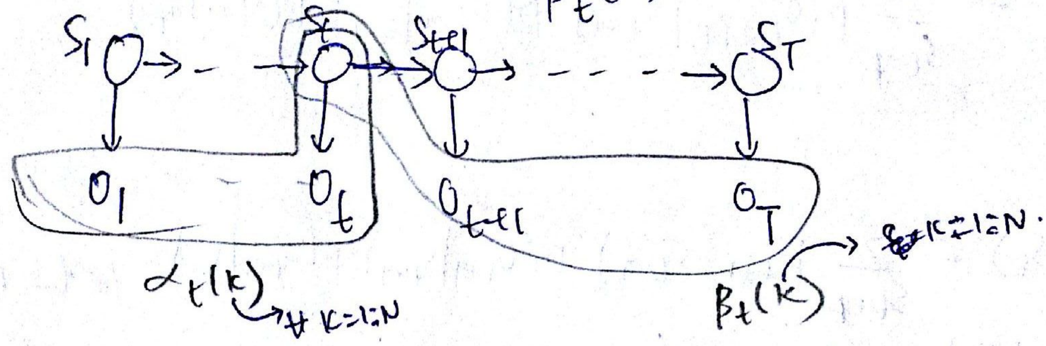
→ Decoding problem - 1 & 2 :-

find $p(s_t=k \mid \{o_t\}_{t=1}^{T}, \lambda)$. λ is the HMM model

$$p(s_t=k, \{o_t\}_{t=1}^{T}) = p(o_1, o_2, \dots o_t, s_t=k, o_{t+1}, \dots o_T)$$

$$= p(o_{t+1}, \dots o_T \mid s_t=k, o_{1:t})\, p(s_t=k, o_{1:t})$$

$$= \underbrace{p(o_{t+1:T} \mid s_t=k)}_{\beta_t(k)} \cdot \underbrace{p(s_t=k, o_{1:t})}_{\alpha_t(k)}$$



$$\alpha_t(k) = p(o_{1:t}, s_t=k) = \sum_{s_{t-1}=1}^{N} p(o_{1:t}, s_t=k, s_{t-1})$$

$$\cancel{\sum_{s_{t-1}} p(s_t \mid s_{t-1}, o_{1:t}) \cdot p(s_{t-1}, o}$$

$$= \sum_{s_{t-1}} p(o_t \mid o_{1:t-1}, s_t, s_{t-1}) \cdot p(o_{1:t-1}, s_t, s_{t-1})$$

$$= \sum_{s_{t-1}} p(o_t \mid s_t) \cdot p(s_t \mid s_{t-1}, o_{1:t-1}) \cdot p(s_{t-1}, o_{1:t-1})$$

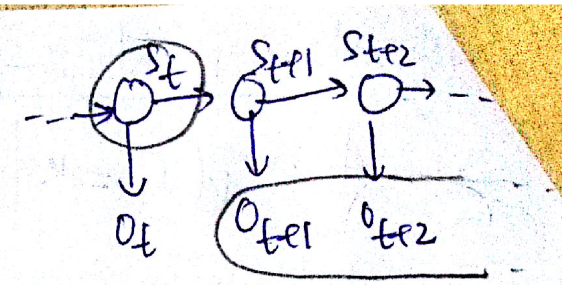
$$= \sum_{s_{t-1}} \underbrace{p(o_t \mid s_t) \cdot p(s_t \mid s_{t-1})}_{\text{known}} \cdot \underbrace{p(s_{t-1}, o_{1:t-1})}_{\alpha_{t-1}(s_{t-1})}$$

$$\Rightarrow \alpha_t(s_t) = \sum_{s_{t-1}=1}^{N} p(o_t \mid s_t)\, p(s_t \mid s_{t-1}) \cdot \alpha_{t-1}(s_{t-1})$$

for $s_{t-1} = 1,2,3,\dots N$.
$t = 2,3,\dots T$.

$$\alpha_1(s_1) = p(s_1, o_1) = p(o_1 \mid s_1) \cdot p(s_1)$$

$\rightarrow \beta_t(k) = p(o_{t+1}, o_{t+2}, \cdots o_T \mid s_t = k)$



$$\beta_t(s_t) = p(o_{t+1}, o_{t+2}, \cdots o_T \mid s_t).$$

$$= \sum_{s_{t+1}}^{N} p(o_{t+1:T}, s_{t+1} \mid s_t)$$

$$= \cancel{\sum_{s_{t+1}} p(o_{t+1} \mid s_{t+1}, o_{t+2:}}$$

$$= \sum_{s_{t+1}} p(o_{t+2:T} \mid s_{t+1}, o_{t+1}, s_t) \cdot p(s_{t+1}, o_{t+1} \mid s_t)$$

$$= \sum_{s_{t+1}} p(o_{t+2:T} \mid s_{t+1}) \cdot p(o_{t+1} \mid s_{t+1}, s_t) \cdot p(s_{t+1} \mid s_t)$$

$$= \sum_{s_{t+1}} \underbrace{p(o_{t+2:T} \mid s_{t+1})}_{\beta_{t+1}(s_{t+1})} \underbrace{p(o_{t+1} \mid s_{t+1}) \cdot p(s_{t+1} \mid s_t)}_{\text{known}}$$

$$\Rightarrow \beta_t(s_t) = \sum_{s_{t+1}} \beta_{t+1}(s_{t+1}) \cdot p(o_{t+1} \mid s_{t+1}) \cdot p(s_{t+1} \mid s_t) \quad \text{for } t = 1, 2, \cdots, T-1.$$

$$\beta_T(s_T) = 1. \quad \forall s_T.$$

***Viterbi Algorithm:-** (proposed in 1967)

Required

> **Remark:**
> if $f(a) \geq 0, \forall a$
> and $g(a,b) \geq 0 \ \forall a, b.$
> then
> $\max\limits_{a,b} f(a) \cdot g(a,b) = \max\limits_{a} \left[f(a) \max\limits_{b} g(a,b) \right]$

find $s^* = \underset{s_1, s_2, \cdots s_T}{\operatorname{argmax}} \; p(\{s_t\}_{t=1}^{T} \mid \{o_t\}_{t=1}^{T})$

$$= \underset{s_t \in \{s_1, s_2, \cdots s_T\}}{\operatorname{argmax}} p(\{s_t\}_{t=1}^{T}, \{o_t\}_{t=1}^{T})$$

$$= \max_{s_1, s_2 \cdots s_T} p(\{s_t\}_{t=1}^{T}, \{o_t\}_{t=1}^{T})$$

$$= \max_{s_1, \cdots s_{t-1}} \max_{s_t} p(s_t, o_t, s_{1:t-1}, o_{1:t-1})$$

$$= \max_{s_1, \cdots s_{t-1}} \max_{s_t} p(o_t \mid s_t) \cdot p(s_t \mid s_{t-1}) \cdot p(s_{1:t-1}, o_{1:t-1})$$

$$= \max_{s_t} \max_{s_1, \cdots s_{t-1}} p(o_t \mid s_t) p(s_t \mid s_{t-1}) \cdot p(s_{1:t-1}, o_{1:t-1})$$

$=$

Generated by CamScanner

Viterbi Algorithm :- (proposed in 1967)

find $s^* = \underset{s_1, s_2 \dots s_T}{\arg\max}\; p(\{s_t\}_{t=1}^{T} \mid \{o_t\}_{t=1}^{T})$.

$$\max_{s_1, s_2 \dots s_t} p(s_{1:t} \mid o_{1:t}) = \max_{s_1, s_2, \dots s_t} p(s_{1:t}, o_{1:t})$$

$$\max_{s_{1:t}} p(s_{1:t} \mid o_{1:t}) = \max_{s_t} \underbrace{\max_{s_1, s_2 \dots s_{t-1}} p(s_{1:t}, o_{1:t})}_{V_t(s_t)}$$

$$\rightarrow V_t(s_t) = \max_{s_{1:t-1}} p(s_t, o_t, s_{1:t-1}, o_{1:t-1})$$

$$= \max_{s_{t-1}} \max_{s_{1:t-2}} \underbrace{p(o_t \mid s_t) \cdot p(s_t \mid s_{t-1})}_{\text{Known}} \cdot p(s_{1:t-1}, o_{1:t-1})$$

$$= \max_{s_{t-1}} \underbrace{p(o_t \mid s_t) \cdot p(s_t \mid s_{t-1})}_{\text{Known}} \cdot \underbrace{\max_{s_{1:t-2}} p(s_{1:t-1}, o_{1:t-1})}_{V_{t-1}(s_{t-1})}$$

$$V_t(s_t) = \max_{s_{t-1}} p(o_t \mid s_t) p(s_t \mid s_{t-1}) \cdot V_{t-1}(s_{t-1})$$

$$\boxed{V_t(s_t) = p(o_t \mid s_t) \cdot \max_{s_{t-1}} p(s_t \mid s_{t-1}) \cdot V_{t-1}(s_{t-1}).} \quad t = 2, 3, \dots T.$$

$$V_1(s_1) = p(s_1, o_1) = p(o_1 \mid s_1) \cdot p(s_1).$$

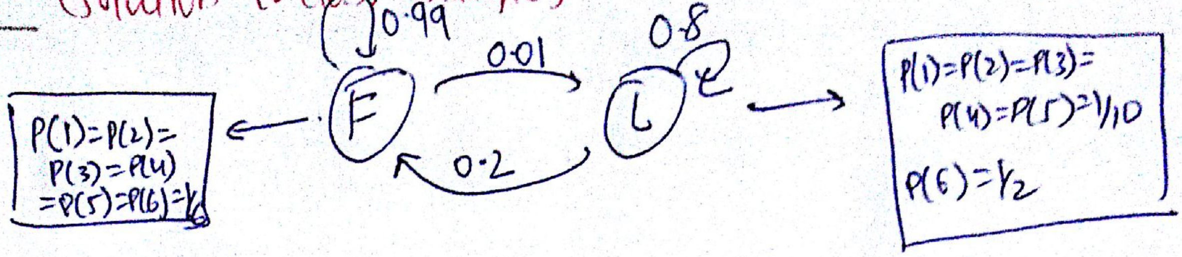
Traceback

$$\rightarrow s_T^* = \underset{s_T}{\arg\max}\; V_T(s_T)$$

$$s_{t-1}^* = \underset{s_{t-1}}{\arg\max}\; p(s_t^* \mid s_{t-1}) V_{t-1}(s_{t-1})$$

Example :- (Solution to class example).

$P(F)=P(L)=1/2$



(1)

Given HMM model and observation seq for 3 timesteps, $O=\{6,2,6\}$. find the most likely state sequence.

	$o_1=6$	$o_2=2$	$o_3=6$
F	$V_1(F) = P(o_1\|F)\cdot P(F)$ $= P(6\|F)\cdot P(F)$ $= \frac{1}{6}\times\frac{1}{2} = \frac{1}{12}$	$V_2(F) = P(o_2\|F)\max\{P(F\|F)\,V_1(F), P(F\|L)\,V_1(L)\}$ $= P(2\|F)\max\{\frac{1}{12}\times 0.99, \frac{1}{4}\times 0.2\}$ $= \frac{1}{6}\times 0.0825 = 0.01375$	$= 0.00226875$
L	$V_1(L) = P(6\|L)\cdot P(L)$ $= \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$	$V_2(L) = P(o_2\|L)\max\{P(L\|F)V_1(F), P(L\|L)\,V_1(L)\}$ $= \frac{1}{10}\max\{0.01\times\frac{1}{12}, 0.8\times\frac{1}{4}\}$ $= 0.02$	$V_3(L) = 0.08$
	$t=1$	$t=2$	$t=3$

$$= \arg\max_{s_1} P(L|s_1)\cdot V_1(s_1)$$

$$= \arg\max_{s_1} \left\{ \begin{array}{l} 0.01\times\frac{1}{12}, \\ 0.8\times\frac{1}{4}, \end{array} \right\}$$

$$= L$$

$$= \arg\max_{s_2} P(s_3|s_2)\, V_2(s_2)$$

$$= \arg\max_{s_2} P(L|s_2)\cdot V_2(s_2)$$

$$= \arg\max_{s_2} \left\{ \begin{array}{l} 0.01\times 0.01375, \\ 0.8\times 0.02 \end{array} \right\}$$

$$= L$$

$$\arg\max_{s_3} V_3(s_3) = L$$

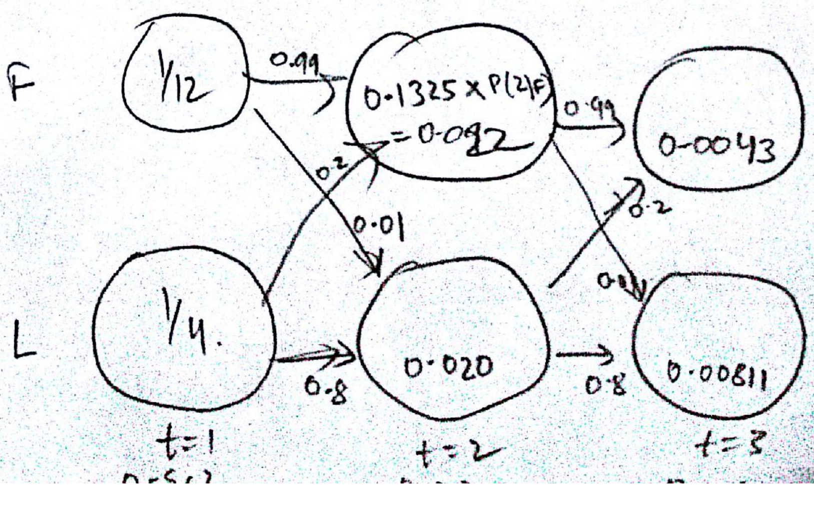
---

(2) $P(\{6,2,6\}) = ?$

$$\alpha_2(F) = \left(\frac{1}{12}\times 0.99 + \frac{1}{4}\times 0.2\right)\times P(2|F)$$

$$= (0.0825 + 0.05)\times\frac{1}{6} = 0.022$$

$$\alpha_2(L) = \left(\frac{1}{12}\times 0.01 + \frac{1}{4}\times 0.8\right)\times P(2|L) = 0.020$$



$$P(\{6,2,6\}) = \sum_{s_t}\alpha_3(s_t) = 0.0043 + 0.0081$$

$$= 0.0124$$