CSE 250A HW 7

Jiping Lin A15058075

$$\begin{tikzcd}
O_1 & O_2 & & O_t & O_{t+1} & & O_T \\
S_1 \arrow[r] \arrow[u] & S_2 \arrow[r] \arrow[u] & \cdots \arrow[r] & S_t \arrow[r] \arrow[u] & S_{t+1} \arrow[r] \arrow[u] & \cdots \arrow[r] & S_T \arrow[u]
\end{tikzcd}$$

7.1 A HOUSE DIVIDED AGAINST ITSELF $\hookrightarrow$ CANNOT STAND

7.2 (a) $P(S_{t+1}=j \mid S_t=i, O_1, \cdots, O_T)$

$$= \frac{P(S_{t+1}=j, S_t=i, O_1, \cdots, O_t, O_{t+1}, \cdots, O_T)}{P(S_t=i, O_1, \cdots, O_t, O_{t+1}, \cdots, O_T)}$$

$P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j, S_t=i, O_1, \cdots, O_{t+1}$

$$= \frac{P(S_t=i, O_1, \cdots, O_t) P(S_{t+1}=j \mid S_t=i, O_1, \cdots, O_t) P(O_{t+1} \mid S_{t+1}=j, S_t=i, O_1, \cdots, O_t) \checkmark}{P(S_t=i, O_1, \cdots, O_t) P(O_{t+1}, \cdots, O_T \mid S_t=i, O_1, \cdots, O_t)}$$

[Product Rule split in 4 pts]

$$= \frac{\alpha_{it} \overset{\text{dsep.(1)}}{P(S_{t+1}=j \mid S_t=i)} \overset{\text{d.sep(1)(2)}}{P(O_{t+1} \mid S_{t+1}=j)} \overset{\text{d-sep(1)}}{P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j)}}{\alpha_{it} \underset{\text{d.sep(1)\&(2) } (O_t \to S_{t+1})}{P(O_{t+1}, \cdots, O_T \mid S_t=i)}}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{it} \beta_{it}}$$

$$= \frac{a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\beta_{it}}$$

(b) $P(S_t=i \mid S_{t+1}=j, O_1, \cdots, O_T)$

$$= \frac{P(S_t=i, S_{t+1}=j, O_1, \cdots, O_T)}{P(S_{t+1}=j, O_1, \cdots, O_{t+1}, O_{t+2}, \cdots, O_T)}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{P(S_{t+1}=j, O_1, \cdots, O_{t+1}) P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j, O_1, \cdots, O_{t+1})} \quad \leftarrow \text{Same as num. in (a)}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{j(t+1)} P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j)} \quad \text{d.sep(1), (2)}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{j(t+1)} \beta_{j(t+1)}}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1})}{\alpha_{j(t+1)}}$$

(c) $P(S_{t-1}=i, S_t=k, S_{t+1}=j \mid O_1, \cdots, O_T)$

$$= \frac{P(S_{t-1}=i, S_t=k, S_{t+1}=j, O_1, \cdots, O_{t-1}, O_t, O_{t+1}, \cdots, O_T)}{P(O_1, \cdots, O_T)}$$

$$= \frac{P(S_{t-1}=i, O_1, \cdots, O_{t-1}) P(S_t=k, S_{t+1}=j, O_t, \cdots, O_T \mid S_{t-1}=i, O_1, \cdots, O_{t-1})}{P(O_1, \cdots, O_T)}$$

NEXT PAGE

$$= \frac{\alpha_{i(t-1)} P(S_t=k \mid S_{t-1}=i, O_1, \cdots O_{t-1}) P(S_{t+1}=j, O_t, \cdots, O_T \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_{t-1})}{P(O_1, \cdots, O_T)}$$

d.sep(1)(2)

$$= \frac{\alpha_{i(t-1)} P(S_t=k \mid S_{t-1}=i) P(O_t \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_{t-1}) P(S_{t+1}=j, O_{t+1}, \cdots, O_T \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_t)}{P(O_1, \cdots, O_T)}$$

d.sep(1)

$$= \frac{\alpha_{i(t-1)} a_{ik} P(O_t \mid S_t=k) P(S_{t+1}=j \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_t) P(O_{t+1}, \cdots, O_T \mid S_{t+1}=j, S_{t-1}=i, S_t=k, O_1, \cdots, O_t)}{P(O_1, \cdots, O_T)}$$

d.sep(1)(2)

$$= \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) P(S_{t+1}=j \mid S_t=k) P(O_{t+1} \mid S_{t+1}=j, S_{t-1}=i, S_t=k, O_1, \cdots, O_t) P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j, S_{t-1}=i, S_t=k, O_1, \cdots, O_{t+1})}{P(O_1, \cdots, O_T)}$$

$$= \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} P(O_{t+1} \mid S_{t+1}=j) P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j)}{P(O_1, \cdots, O_T)}$$

$$= \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)}}{\sum_x P(O_1, \cdots, O_T, S_t=x)} \quad (*)$$

$$= (*) / \sum_x P(S_t=x, O_1, \cdots, O_t) P(O_{t+1}, \cdots, O_T \mid S_t=x, O_1, \cdots, O_t)$$

$$= (*) / \sum_x \alpha_{xt} P(O_{t+1}, \cdots, O_T \mid S_t=x)$$

$$= \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)}}{\sum_x \alpha_{xt} \beta_{xt}}$$

(d) $P(S_{t+1}=j \mid S_{t-1}=i, O_1, \cdots, O_T)$

$$= \frac{P(S_{t+1}=j, S_{t-1}=i, O_1, \cdots, O_T)}{P(S_{t-1}=i, O_1, \cdots, O_T)}$$

$$= \frac{\sum_k P(S_{t+1}=j, S_t=k, S_{t-1}=i, O_1, \cdots, O_T)}{P(S_{t-1}=i, O_1, \cdots, O_{t-1}) P(O_t, \cdots, O_T \mid S_{t-1}=i, O_1, \cdots, O_{t-1})}$$

(c) num.

$$= \frac{\sum_k \alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{i(t-1)} P(O_t, \cdots, O_T \mid S_{t-1}=i)}$$

d.sep(1)

$$= \frac{\sum_k \alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{i(t-1)} \beta_{i(t-1)}}$$

7.3 FTFFTFTTFFTT

7.4 (a) $q_{jt} = P(S_t=j \mid O_1,\cdots,O_{t-1},O_t)$

$$= \frac{P(S_t=j \mid O_1,\cdots,O_{t-1})\,P(O_t \mid S_t=j,O_1,\cdots,O_{t-1})}{P(O_t \mid O_1,\cdots,O_{t-1})}$$ 「Baye's Rule」

$$= \frac{P(S_t=j \mid O_1,\cdots,O_{t-1})\,P(O_t \mid S_t=j)}{P(O_t \mid O_1,\cdots,O_{t-1})}$$ ← d-sep(1)

← denom.

$$= \sum_i P(S_{t-1}=i, S_t=j \mid O_1,\cdots,O_{t-1})\, b_j(O_t) \,/\, \text{denom.}$$

$$= \sum_i P(S_{t-1}=i \mid O_1,\cdots,O_{t-1})\, P(S_t=j \mid S_{t-1}=i, O_1,\cdots,O_{t-1})\, b_j(O_t) \,/\, \text{denom.}$$

$$= b_j(O_t)\sum_i q_{i(t-1)}\, P(S_t=j \mid S_{t-1}=i) \,/\, \text{denom}$$ ← d-sep(1)(2)

$$= b_j(O_t)\sum_i q_{i(t-1)}\, a_{ij}$$

denom $= P(O_t \mid O_1,\cdots,O_{t-1}) = \sum_j P(S_t=j, O_t \mid O_1,\cdots,O_{t-1})$

$$= \sum_j P(S_t=j \mid O_1,\cdots,O_{t-1})\, P(O_t \mid S_t=j, O_1,\cdots,O_{t-1})$$

$$= \sum_j b_j(O_t) \sum_i q_{i(t-1)}\, a_{ij} \quad \text{(from num.)}$$

So $$q_{jt} = \frac{b_j(O_t)\sum_i q_{i(t-1)} a_{ij}}{\sum_{ij} b_j(O_t)\, q_{i(t-1)}\, a_{ij}}$$

(b) $$P(X_t \mid Y_1,\cdots,Y_t) = \frac{P(X_t \mid Y_1,\cdots,Y_{t-1})\, P(Y_t \mid X_t, Y_1,\cdots,Y_{t-1})}{P(Y_t \mid Y_1,\cdots,Y_{t-1})}$$

Num. $= P(X_t \mid Y_1,\cdots,Y_{t-1})\, P(Y_t \mid X_t)$ ← d-sep(1)

$$= P(Y_t \mid X_t) \int P(X_t, X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1} \quad \text{(Marg. on } X_{t-1})$$

$$= P(Y_t \mid X_t) \int P(X_t \mid X_{t-1}, Y_1,\cdots,Y_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}$$

$$= P(Y_t \mid X_t) \int \underset{\text{d-sep(1)}}{P(X_t \mid X_{t-1})}\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}$$

Denom $= \int P(X_t, Y_t \mid Y_1,\cdots,Y_{t-1})\, dX_t$

$$= \int \underbrace{P(X_t \mid Y_1,\cdots,Y_{t-1})\, P(Y_t \mid X_t, Y_1,\cdots,Y_{t-1})}_{\text{num.}}\, dX_t$$

$$= \int P(Y_t \mid X_t)\left(\int P(X_t \mid X_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}\right) dX_t .$$

it is easier than other distributions where we have to compute each cond. prob. and find integral.

So $$P(X_t \mid Y_1,\cdots,Y_t) = \text{Num./denom.} = \frac{P(Y_t \mid X_t)\int P(X_t \mid X_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}}{\int P(Y_t \mid X_t)\left(\int P(X_t \mid X_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}\right) dX_t .}$$

If X, Y are Gaussian, then $Y \mid X$ is also gaussian, then since the integral of Gaussian is easy to compute 」

7.5 (a) $P(Y_1=j, O_1=o_1)$

$= \sum_i P(Y_1=j, X_1=i, O_1=o_1)$

$= \sum_i P(Y_1=j) P(X_1=i \mid Y_1=j) P(O_1=o_1 \mid X_1=i, Y_1=j)$

$= \sum_i P(X_1=i, Y_1=j) b_{ij}(o_1)$

$= \sum_i P(X_1=i) P(Y_1=j) b_{ij}(o_1)$

↳ d-sep(3)

$= \sum_i P(X_1=i) \pi_j b_{ij}(o_1)$

(b) Assume $t>1$,

$\alpha_{jt} = P(o_1, \cdots, o_t, Y_t=j)$

$= \sum_k \sum_i P(Y_{t-1}=k, Y_t=j, X_t=i, o_1, \cdots, o_t)$

$= \sum_k \sum_i P(o_1, \cdots, o_{t-1}, Y_{t-1}=k) P(X_t=i \mid Y_{t-1}=k, o_1, \cdots, o_{t-1})$

$P(Y_t=j \mid X_t=i, Y_{t-1}=k, o_1, \cdots, o_{t-1})$

$P(o_t \mid X_t=i, Y_{t-1}=k, Y_t=j, o_1, \cdots, o_{t-1})$

$= \sum_k \sum_i \alpha_{k(t-1)} \underbrace{P(X_t=i \mid Y_{t-1}=k)}_{\text{d-sep(2)}} \underbrace{P(Y_t=j)}_{\text{d-sep(3)}} \underbrace{P(o_t \mid X_t=i, Y_t=j)}_{\text{d-sep(1)}}$

$= \sum_k \sum_i \alpha_{k(t-1)} a_{ki} \pi_j b_{ij}(o_t)$

↑ recursion

where $t=1$ case is part (a)

(c) $P(o_1, \cdots, o_T) = \sum_x P(o_1, \cdots, o_T, Y_T=j) = \sum_x \alpha_{xT}$

(d) $P(o_1, \cdots, o_T) = \sum_x \alpha_{xT} = \sum_x \sum_k \sum_i \alpha_{k(T-1)} a_{ki} \pi_x b_{ix}(o_T)$

α_{k1} takes $O(n_x)$ since sum is from $i=1$ to n_x

$\alpha_{kT} = O(n_y O(\alpha_{k(T-1)})) \Rightarrow \alpha_{kT}$ takes $O((T-1) n_x n_y)$

↳ sum from $k=1$ to n_y, recurse from T to 1

then $\sum_x \alpha_{xt}$ takes $n_y O(\alpha_{xt})$ so the final complexity is $O(n_x n_y^2 (T-1))$