CS221 Fall 2018 - 2019 Homework 7

Name: Dat Nguyen
Date: 4/8/2019

By turning in this assignment, I agree by the Stanford honor code and declare that all of this is my own work.

Problem 1: Warmup

(a) We have

$$\begin{aligned}
&\mathbb{P}(C_2 = 1 | D_2 = 0) \\
&\propto \mathbb{P}(C_2 = 1, D_2 = 0) \\
&= \sum_{C_1, D_1, C_3, D_3} \mathbb{P}(C_1)\mathbb{P}(D_1|C_1)\mathbb{P}(C_2 = 1|C_1)\mathbb{P}(D_2 = 0|C_2 = 1)\mathbb{P}(C_3|C_2 = 1)\mathbb{P}(D_3|C_3) \\
&= \sum_{C_1} \mathbb{P}(C_1)\mathbb{P}(C_2 = 1|C_1)\eta \\
&= 0.5 \times \epsilon\eta + 0.5 \times (1-\epsilon)\eta
\end{aligned}$$

Similarly we get

$$\mathbb{P}(C_2 = 0 | D_2 = 0) \propto 0.5 \times \epsilon(1-\eta) + 0.5 \times (1-\epsilon)(1-\eta)$$

Normalizing

$$\begin{aligned}
\mathbb{P}(C_2 = 1 | D_2 = 0) &= \frac{\epsilon\eta + (1-\epsilon)\eta}{\epsilon\eta + (1-\epsilon)\eta + \epsilon(1-\eta) + (1-\epsilon)(1-\eta)} \\
&= \eta
\end{aligned}$$

(b) We have

$$\begin{aligned}
&\mathbb{P}(C_2 = 1 | D_2 = 0, D_3 = 1) \\
&\propto \mathbb{P}(C_2 = 1, D_2 = 0, D_3 = 1) \\
&= \sum_{C_1, D_1, C_3} \mathbb{P}(C_1)\mathbb{P}(D_1|C_1)\mathbb{P}(C_2 = 1|C_1)\mathbb{P}(D_2 = 0|C_2 = 1)\mathbb{P}(C_3|C_2 = 1)\mathbb{P}(D_3 = 1|C_3) \\
&= \sum_{C_1} \mathbb{P}(C_1)\mathbb{P}(C_2 = 1|C_1)\mathbb{P}(D_2 = 0|C_2 = 1) \sum_{C_3} \mathbb{P}(C_3|C_2 = 1)\mathbb{P}(D_3 = 1|C_3) \\
&= \mathbb{P}(D_2 = 0|C_2 = 1) \sum_{C_1} \mathbb{P}(C_1)\mathbb{P}(C_2 = 1|C_1) \sum_{C_3} \mathbb{P}(C_3|C_2 = 1)\mathbb{P}(D_3 = 1|C_3) \\
&= \eta \times 0.5 \times ((1-\epsilon)(1-\eta) + \epsilon\eta)
\end{aligned}$$

Similary we get

$$\mathbb{P}(C_2 = 0|D_2 = 0, D_3 = 1) \propto (1-\eta) \times 0.5 \times (\epsilon(1-\eta) + (1-\epsilon)\eta)$$

Normalizing

$$\begin{aligned}
\mathbb{P}(C_2 = 1|D_2 = 0, D_3 = 1) &= \frac{\eta((1-\epsilon)(1-\eta)+\epsilon\eta)}{\eta((1-\epsilon)(1-\eta)+\epsilon\eta) + (1-\eta)(\epsilon(1-\eta)+(1-\epsilon)\eta)} \\
&= \frac{\eta - \eta^2 - \eta\epsilon + 2\eta^2\epsilon}{\epsilon + 2\eta - 4\eta\epsilon - 2\eta^2 + 4\epsilon\eta^2}
\end{aligned}$$

(c) i

$$\begin{aligned}
\mathbb{P}(C_2 = 1|D_2 = 0) &= 0.2 \\
\mathbb{P}(C_2 = 1|D_2 = 0, D_3 = 1) &\approx 0.416
\end{aligned}$$

ii Adding second sensor reading $D_3 = 1$ increases our belief about $C_2 = 1$. This coincides with our intuition that $D_3 = 1$ makes it more likely for $C_3 = 1$ which in turn makes it more likely for $C_2 = 1$.

iii From equations for $\mathbb{P}(C_2 = 1|D_2 = 0) = 0.2$ and $\mathbb{P}(C_2 = 1|D_2 = 0, D_3 = 1)$ we have

$$\begin{aligned}
\frac{\eta - \eta^2 - \eta\epsilon + 2\eta^2\epsilon}{\epsilon + 2\eta - 4\eta\epsilon - 2\eta^2 + 4\epsilon\eta^2} &= \eta \\
\frac{0.16 - 0.2\epsilon + 0.08\epsilon}{\epsilon + 0.4 - 0.8\epsilon - 0.08 + 0.16\epsilon} &= 0.2 \\
\frac{0.16 - 0.12\epsilon}{0.36\epsilon + 0.32} &= 0.2 \\
0.16 - 0.12\epsilon &= 0.072\epsilon + 0.064 \\
\epsilon &= 0.5
\end{aligned}$$

Although observing $D_3 = 1$ increases our belief about $C_3 = 1$ but since $\epsilon = 0.5$, $C_3 = 1$ does not necessarily make it more likely that $C_2 = 1$ has caused that.

Problem 5: Which car is it?

(a)

$$
\begin{aligned}
&\mathbb{P}(C_{11}, C_{12} | E_1 = e_1) \\
&\propto \mathbb{P}(C_{11}, C_{12}, E_1 = e_1) \\
&\propto \mathbb{P}(C_{11}, C_{12}, (E_1, E_2) = (e_{11}, e_{12})) \\
&\propto 0.5 \times \mathbb{P}\Big(C_{11}, C_{12}, (E_1, E_2) = (d_{11}, d_{12})\Big) \\
&\quad + 0.5 \times \mathbb{P}\Big(C_{11}, C_{12}, (E_{11}, E_{12}) = (d_{12}, d_{11})\Big) \\
&\propto \mathbb{P}\Big(C_{11}, C_{12} | (E_{11}, E_{12}) = (d_{11}, d_{12})\Big) \\
&\quad + \mathbb{P}\Big(C_{11}, C_{12} | (E_{11}, E_{12}) = (d_{12}, d_{11})\Big) \\
&\propto \mathbb{P}\Big(C_{11} | E_{11} = d_{11}\Big)\mathbb{P}\Big(C_{12} | E_{12} = d_{12}\Big) \\
&\quad + \mathbb{P}\Big(C_{11} | E_{12} = d_{11}\Big)\mathbb{P}\Big(C_{12} | E_{11} = d_{12}\Big) \\
&\propto p(C_{11}) p_{\mathcal{N}}(E_{11}; ||e_{11} - c_{11}||, \sigma^2) p(C_{12}) p_{\mathcal{N}}(E_{12}; ||e_{12} - c_{12}||, \sigma^2) \\
&\quad + p(C_{11}) p_{\mathcal{N}}(E_{12}; ||e_{12} - c_{11}||, \sigma^2) p(C_{12}) p_{\mathcal{N}}(E_{11}; ||e_{11} - c_{12}||, \sigma^2)
\end{aligned}
$$

(b) Let the optimal assignment for K cars be $(o_{11}, o_{12}, \dots, o_{1K})$. Using the same argument for K=2 above we arrive at the ditribution over K cars given the sensor reading

$$
\begin{aligned}
&\mathbb{P}(C_{11} = o_{11}, \dots, C_{1K} = o_{1K} | E_1 = e_1) \\
&\propto \sum_{e \in \text{Permutation}(e_1)} \prod_{k=1}^{K} p(o_{1k}) p_{\mathcal{N}}(e_{1k}; ||e_{1k} - o_{1k}) \\
&\propto \sum_{e \in \text{Permutation}(e_1)} p(o_{11})^K \prod_{k=1}^{K} p_{\mathcal{N}}(e_{1k}; ||e_{1k} - o_{1k})
\end{aligned}
$$

Therefore for any permutation of $(o_{11}, o_{12}, \dots, o_{1K})$ we have the same distribution, which shows that there is at least $K!$ number of assignments that obtain the maximum value of that distribution.

(c)

(d) Forward message

Let $p(c_{ti} = j | e_1, \dots, e_t) = \alpha_{ti}(j)$. We have

$$\begin{aligned}
p(c_{ti}|e_1,\dots,e_t) &\propto \sum_{c_{t-1,i}} p(c_{ti}, c_{t-1,i}, e_1,\dots,e_t) \\
&\propto \sum_{c_{t-1,i}} p(c_{t-1,i}, e_1,\dots,e_{t-1}) p(c_{ti}|c_{t-1,i}) p(e_t|c_{ti}, c_{t-1,i}, e_1,\dots,e_{t-1}) \\
&\propto \sum_{c_{t-1,i}=a} \alpha_{t-1,i}(a) p(c_{ti}|c_{t-1,i}) \sum_{m=1}^{K} \frac{1}{K} \Big(p(e_{tm}|O=m, c_{ti}) \\
&\qquad\qquad p(e_{t\backslash m}|O=m, c_{ti}, c_{t-1,i}, e_1,\dots,e_{t-1}) \Big) \\
&\propto \sum_{c_{t-1,i}=a} \alpha_{t-1,i}(a) p(c_{ti}|c_{t-1,i}) \sum_{m=1}^{K} \mathbb{1}(e_{tm} = c_{ti}) \times L \\
&\propto \sum_{c_{t-1,i}=a} \alpha_{t-1,i}(a) p(c_{ti}|c_{t-1,i}) \sum_{m=1}^{K} \mathbb{1}(e_{tm} = c_{ti})
\end{aligned}$$

And

$$\begin{aligned}
p(c_{1i}|e_1) &\propto p(c_{1i}, e_1) \\
&\propto \sum_{m=1}^{K} p(c_{1i}, e_1, O=m) \\
&\propto \sum_{m=1}^{K} p(e_1|c_{1i}, O=m) p(c_{1i}, O=m) \\
&\propto \sum_{m=1}^{K} \mathbb{1}(e_{1m} = c_{1i}) p(O=m|c_{1i}) p(c_{1i}) \\
&\propto p(c_{1i}) \sum_{m=1}^{K} \mathbb{1}(e_{1m} = c_{1i})
\end{aligned}$$

Backward message

Let $p(e_{t+1:T}|c_{ti}=j, e_{1:t}) = \beta_j(t+1)$. We have

$$
\begin{aligned}
p(e_{t+1:T}|c_{ti}, e_{1:t}) &= \sum_{c_{t+1,i}} p(e_{t+1:T}, c_{t+1,i}|c_{ti}, e_{1:t}) \\
&= \sum_{c_{t+1,i}} p(e_{t+1}, e_{t+2:T}, c_{t+1,i}|c_{ti}, e_{1:t}) \\
&= \sum_{c_{t+1,i}} p(e_{t+2:T}|e_{t+1}, c_{t+1,i}, c_{ti}, e_{1:t}) p(e_{t+1}|c_{t+1,i}, c_{ti}, e_{1:t}) p(c_{t+1,i}|c_{ti}, e_{1:t}) \\
&= \sum_{c_{t+1,i}} p(e_{t+2:T}|e_{1:t+1}, c_{t+1,i}) p(c_{t+1,i}|c_{ti}) p(e_{t+1}|c_{t+1,i}, e_{1:t}) \\
&= \sum_{c_{t+1,i}} \beta(t+1) p(c_{t+1,i}|c_{ti}) \sum_{m=1}^{K} p(e_{t+1}, O=m|c_{t+1,i}, e_{1:t}) \\
&= \sum_{c_{t+1,i}} \beta(t+1) p(c_{t+1,i}|c_{ti}) \sum_{m=1}^{K} \frac{1}{K} p(e_{t+1,m}|O=m, c_{t+1,i}, e_{t+1,\backslash m}, e_{1:t}) \\
&\qquad p(e_{t+1,\backslash m}|O=m, c_{t+1,i}, e_{1:t}) \\
&\propto \sum_{c_{t+1,i}} \beta(t+1) p(c_{t+1,i}|c_{ti}) \sum_{m=1}^{K} \mathbb{1}(e_{t+1,m} = c_{t+1,i}) \times L \\
&\propto \sum_{c_{t+1,i}} \beta(t+1) p(c_{t+1,i}|c_{ti}) \sum_{m=1}^{K} \mathbb{1}(e_{t+1,m} = c_{t+1,i})
\end{aligned}
$$

And we let $\beta_j(T+1) = 1$

After calculating the forward and backward message for c_{ti}, we multiply them together to arrive at

$$
p(c_{ti}|e_{1:t}) p(e_{t+1:T}|c_{ti}, e_{1:t}) = p(e_{t+1:T}, c_{ti}|e_{1:t}) \propto p(c_{ti}|e_{1:T})
$$