CS 221 HW - Car

Name: Vinod Kumar Senthil Kumar

Sunet ID: vinodkum

Contributors: Joe Fan

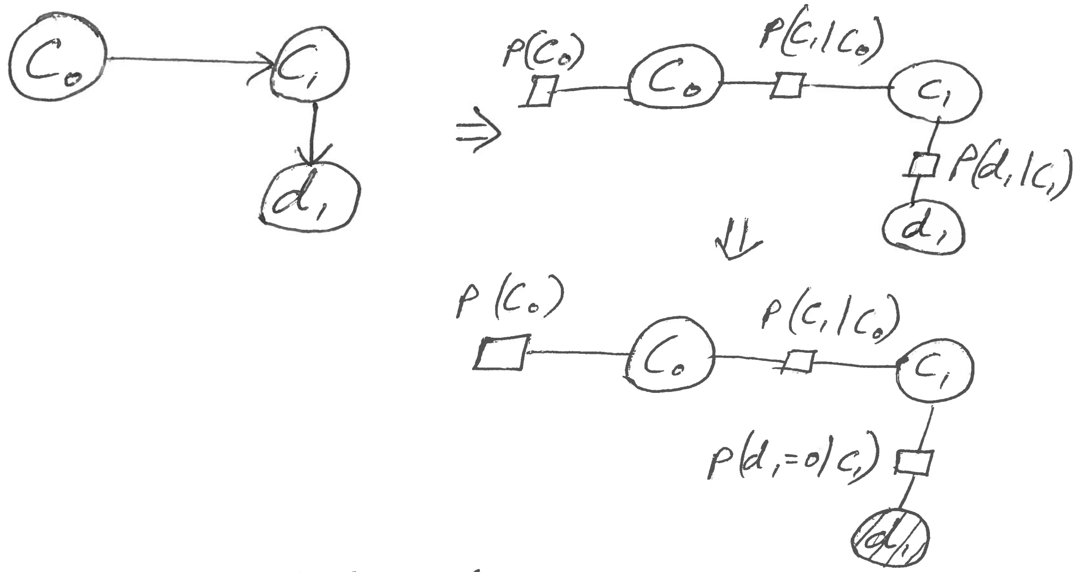
Problem 1:

$$p(c_t \mid c_{t-1}) = \begin{cases} \epsilon & , \text{if } c_t \neq c_{t-1} \\ 1-\epsilon & , \text{if } c_t = c_{t-1} \end{cases}$$

$$p(d_t \mid c_t) = \begin{cases} \eta & , \text{if } d_t \neq c_t \\ 1-\eta & , \text{if } d_t = c_t \end{cases}$$

$$c_t \in \{0,1\} \quad , \quad d_t \in \{0,1\}$$

1a)



$$p(C_1 \mid d_1=0) = \sum_{C_0 \in \{0,1\}} p(C_0)\, p(C_1 \mid C_0)\, p(d_1=0 \mid C_1)$$

$$p(C_1=0 \mid d_1=0) = p(C_0=0)\, p(C_1=0 \mid C_0=0)\, p(d_1=0 \mid C_1=0) + p(C_0=1)\, p(C_1=0 \mid C_0=1)\, p(d_1=0 \mid C_1=0)$$

$$= (0.5)\left[(1-\epsilon)(1-\eta) + \epsilon(1-\eta) \right]$$

$$P(C_1=1 \mid d_1=0) = P(C_0=0)\,P(C_1=1 \mid C_0=0)\,P(d_1=0 \mid C_1=1) + P(C_0=1)\,P(C_1=1 \mid C_0=1)\,P(d_1=0 \mid C_1=1)$$

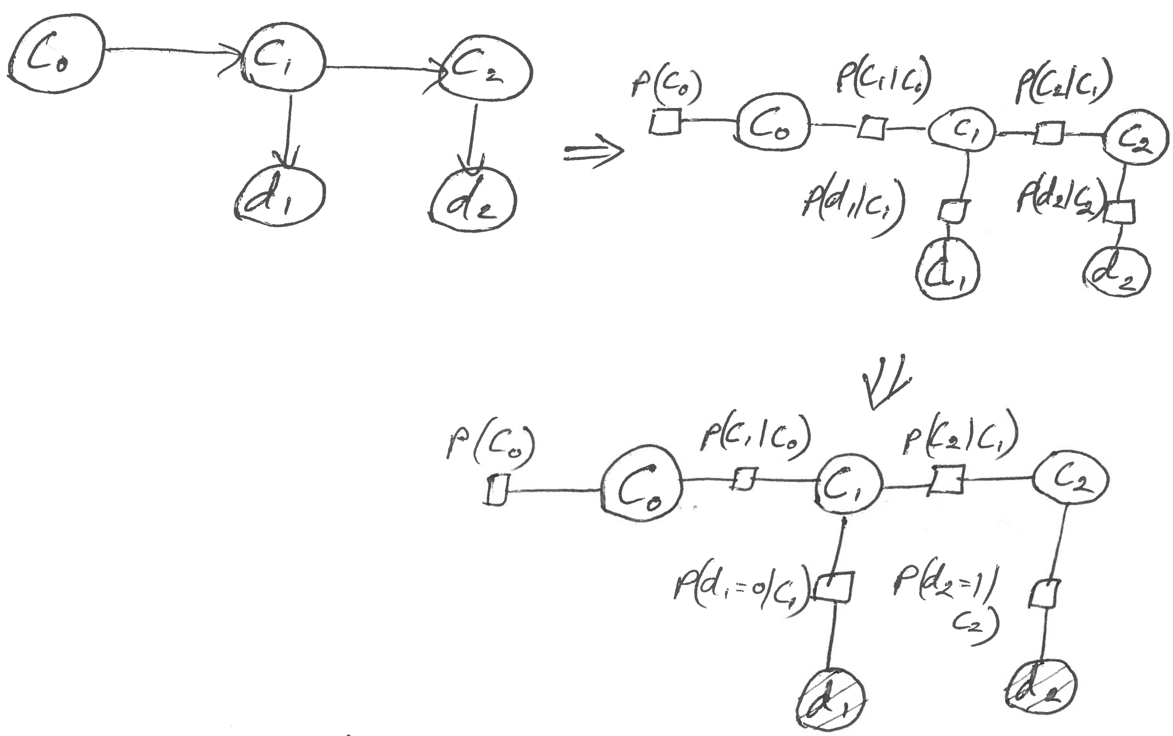
$$= 0.5\left[\epsilon\eta + (1-\epsilon)\eta\right]$$

After normalization,

$$P(C_1=1 \mid d_1=0) = \frac{(0.5)\left[\cancel{\epsilon\eta} + \eta - \cancel{\epsilon\eta}\right]}{(\cancel{0.5})\left(\left[1-\eta-\cancel{\epsilon}+\cancel{\epsilon\eta}+\cancel{\epsilon}-\cancel{\epsilon\eta}\right] + \left[\cancel{\epsilon\eta}+\eta-\cancel{\epsilon\eta}\right]\right)}$$

$$= \frac{\eta}{\left[1-\eta+\eta\right]} = \eta$$

1b)



$$p(c_1 \mid d_1=0, d_2=1) = \sum_{\substack{c_0 \\ \in\{0,1\}}} \sum_{\substack{c_1 \\ \in\{0,1\}}} p(c_0)\, p(c_1|c_0)\, p(c_2|c_1)\, p(d_1=0|c_1)\, p(d_2=1|c_2)$$

$$= \sum_{c_0 \in\{0,1\}} p(c_0)\, p(c_1|c_0)\, p(d_1=0|c_1) \cdot \sum_{c_2\in\{0,1\}} p(c_2|c_1)\, p(d_2=1|c_2)$$

$$p(c_1=0 \mid d_1=0, d_2=1) = \sum_{c_0=\{0,1\}} p(c_0)\, p(c_1=0|c_0)\, p(d_1=0|c_1=0) \cdot \sum_{c_2\in\{0,1\}} p(c_2|c_1=0)\, p(d_2=1|c_2)$$

~~for~~ from 2a, we know

$$p(C_1=0 \mid d_1=0) = \sum_{C_0=\{0,1\}} p(C_0)\, p(C_1=0 \mid C_0)\, p(d_1=0 \mid C_1=0)$$

$$= (0.5)\left[1-\eta \cancel{-\epsilon} \cancel{+\epsilon\eta} \cancel{+\epsilon} \cancel{-\epsilon\eta}\right]$$

$$= (0.5)(1-\eta)$$

Thus, $p(C_1=0 \mid d_1=0, d_2=1) = (0.5)(1-\eta)\left[p(C_2=0 \mid C_1=0) \cdot p(d_2=1 \mid C_2=0) + p(C_2=1 \mid C_1=0) \cdot p(d_2=1 \mid C_2=1)\right]$

$$= (0.5)(1-\eta)\left[(1-\epsilon)\eta + \epsilon(1-\eta)\right]$$

$$= (0.5)(1-\eta)\left[\eta - \epsilon\eta + \epsilon - \epsilon\eta\right]$$

$$= (0.5)(1-\eta)\left[\eta + \epsilon - 2\epsilon\eta\right]$$

$|||^{ly}$ (from 1a) $p(C_1=1 \mid d_1=0, d_2=1) = (0.5)(\eta)\left[p(C_2=0 \mid C_1=1) \cdot p(d_2=1 \mid C_2=0) + p(C_2=1 \mid C_1=1)\, p(d_2=1 \mid C_2=1)\right]$

$$= (0.5)\,\eta\left[(1-\epsilon)(1-\eta) + \epsilon\eta\right]$$

$$= (0.5)\,\eta\left[1-\epsilon-\eta+2\epsilon\eta\right]$$

After normalizing,

$$P(C_1=1 \mid d_1=0, d_2=1) = \frac{(0.5)\,\eta\,[1-\epsilon-\eta+2\epsilon\eta]}{(0.5)\left[\eta(1-\epsilon-\eta+2\epsilon\eta)+(1-\eta)(\epsilon+\eta-2\epsilon\eta)\right]}$$

$$= \frac{\eta-\epsilon\eta-\eta^2+2\epsilon\eta^2}{\eta-\epsilon\eta-\eta^2+2\epsilon\eta^2+\epsilon+\eta-2\epsilon\eta-\epsilon\eta-\eta^2+2\epsilon\eta^2}$$

$$= \frac{\eta-\epsilon\eta-\eta^2+2\epsilon\eta^2}{2(\eta-2\epsilon\eta-\eta^2+2\epsilon\eta^2)+\epsilon}$$

1c) Substituting $\epsilon = 0.1$, $\eta = 0.2$, we get

$$P(C_1=1 \mid d_1=0) = 0.2$$
$$P(C_1=1 \mid d_1=0, d_2=1) = 0.4157$$

Adding $d_2=1$ increased the probability from 0.2 to 0.4157. Intutively, this makes sense as more consistent observation will increase probability.

After adding the second reading $d_2 = 1$, the probability of $C_1 = 1$ has increased because the car has 90% ~~of~~ chance of staying in the same location, and the sensor is 80% accurate. Intuitively, it requires some "effort" for the car to move from 0 to 1.

Problem 5:

5 a) $$p(C_{11}, C_{12} | e_1) = p(C_{11}, C_{12}) \cdot p(e_1 | C_{11}, C_{12})$$

(By Bayes' theorem)

~~pe~~ e_1 is a collection $\{e_{11}, e_{12}\}$ where $\{C_{11}, C_{12}\}$ can correspond to $\{e_{11}, e_{12}\}$ or $\{e_{12}, e_{11}\}$ with equal probability.

$$p(e_1 | C_{11}, C_{12}) = (0.5)\left[p(e_{11}|C_{11}) \cdot p(e_{12}|C_{12}) + p(e_{12}|C_{11}) \cdot p(e_{11}|C_{12})\right]$$

where $p(e_{1i} | C_{1i}) = P_N(e_{1i}; \|a_1 - C_{1i}\|, \sigma^2)$

Thus,

$$p(C_{11}, C_{12} | e_1) = p(C_{11}, C_{12})\left(\frac{1}{2}\right)\left[P_N(e_{11}; \|a_1 - C_{11}\|, \sigma^2) \cdot P_N(e_{12}; \|a_1 - C_{12}\|, \sigma^2) + P_N(e_{12}; \|a_1 - C_{11}\|, \sigma^2) \cdot P_N(e_{11}; \|a_1 - C_{12}\|, \sigma^2)\right]$$

5b) Let $(C_{11}, \ldots, C_{1K}) = (V_{11}, \ldots, V_{1K})$ be an assignment that produces the maximum value of $P(C_{11}, \ldots, C_{1K} | e_1)$.

WKT, $P(C_{11}, \ldots, C_{1K} | e_1) = P(C_{11}, \ldots, C_{1K}) \left(\frac{1}{K!}\right)$

$$\Big[P(e_{11}|C_{11}) \cdot P(e_{12}|C_{12}) \cdots P(e_{1K}|C_{1K}) +$$
$$P(e_{12}|C_{11})\, P(e_{11}|C_{12}) \cdots P(e_{1K}|C_{1K}) +$$
$$\vdots$$
$$\text{all } K! \text{ permutations} \Big]$$

Where $P(e_{1i}|C_{1j}) = P_N(e_{1i}; \|a_1 - c_{1j}\|, \sigma^2)$

Assuming the priors $P(C_{1i})$ are the same for all i,

$$P(C_{11}, \ldots, C_{1K}) = \prod_{i=1}^{K} \cancel{P(C_{1K})}\, P(C_{1i})$$

$$P(C_{11}, \ldots, C_{1K} | e_1) \propto \Big[P(e_{11}|C_{11})\, P(e_{12}|C_{12}) \cdots P(e_{1K}|C_{1K}) +$$
$$\vdots$$
$$K! \text{ permutations} \Big]$$

~~as~~ for different value permutation of $C_{11}, \ldots, C_{1K}$ over $V_1, \ldots, V_{1K}$ as

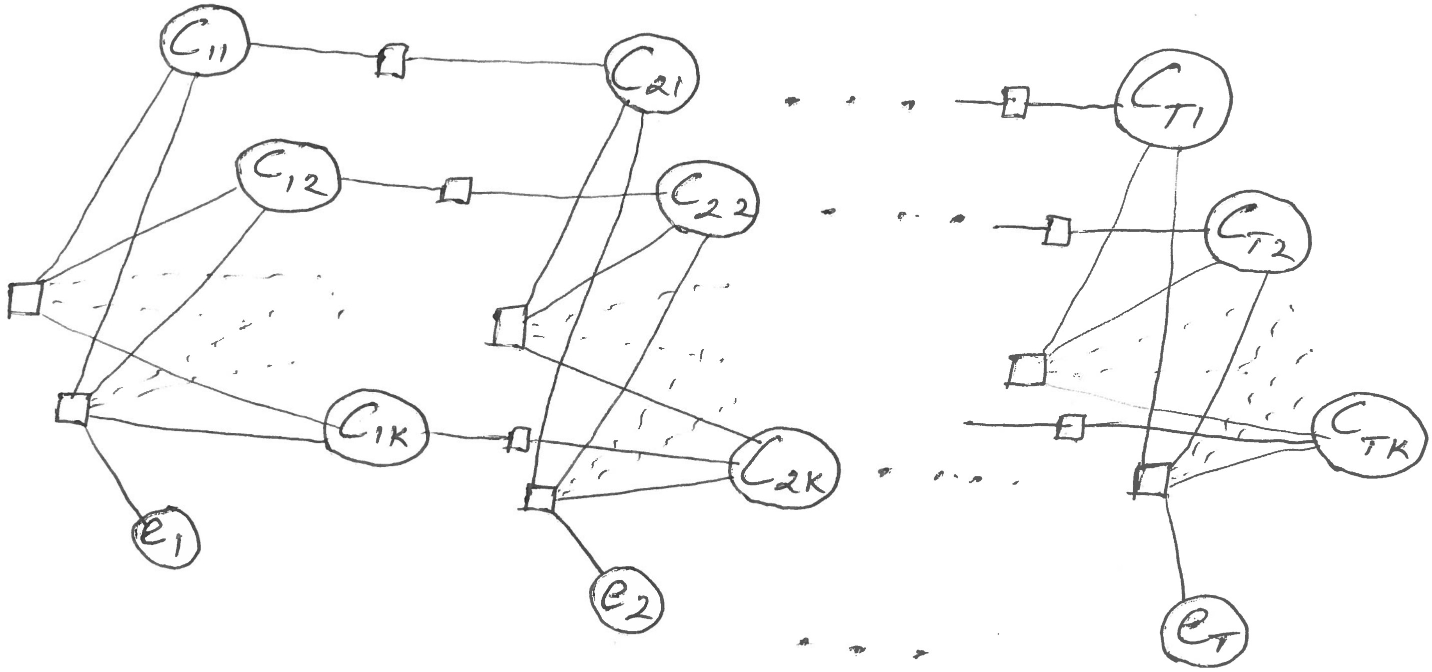
the joint prior $P(C_{11}, \ldots, C_{1K})$ is the same.

On observing the sum of the $K!$ permutations of $P(e_{1i} | c_{1j})$ for $i, j \in \{1, \dots, K\}$, we can see that the sum will be the same for all permutation assigments of $(c_{11}, \dots, c_{1K})$ over $(v_{11}, \dots, v_{1K})$.

As $(c_{11}, \dots c_{1K}) = (v_{11}, \dots, v_{1K})$ produces the maximum value of $P(c_{11}, \dots, c_{1K} | e_1)$, all permutation assignments of $(c_{11}, \dots, c_{1K})$ over $(v_{11}, \dots, v_{1K})$ gives the ~~m~~ maximum probability value.

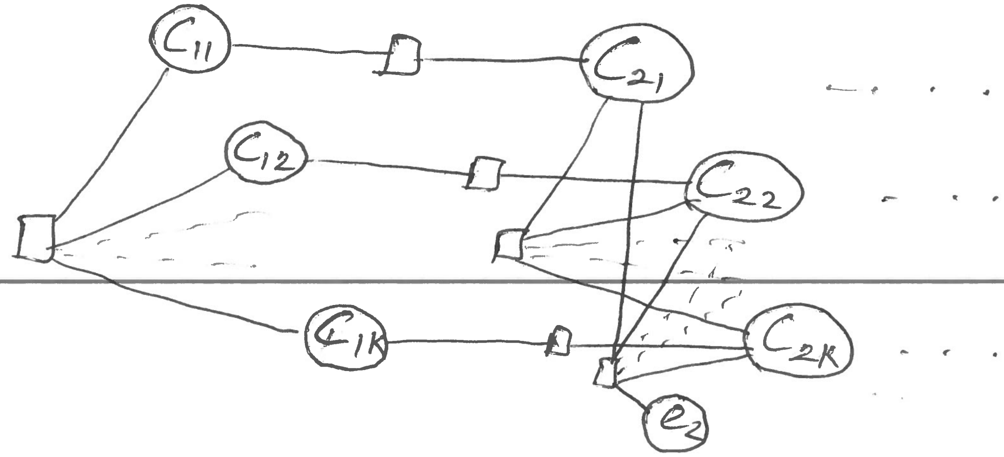
Thus there are atleast $K!$ values of $(c_{11}, \dots, c_{1K})$ that gives the maximum value of $P(c_{11}, \dots, c_{1K} | e_1)$.

5c) The factor graph would be as follows:

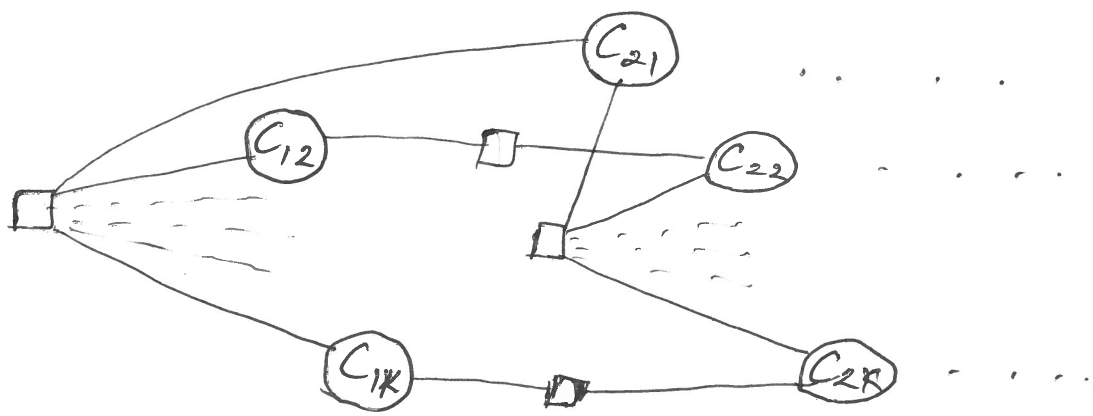


The evidence variables have the smallest Markov blanket. Removing them would introduce a K-nary potential connecting $C_{t1}, \dots C_{tK}$ for e_t.

Eg: if we remove e_1,



After removing all evidence variables, we remove the car position variables. Say we remove C_{11}. This introduces a K-nary potential connecting its Markov Blanket of $C_{12}, C_{13}, \ldots C_{1K}$ & C_{21}.



Similarly, removing C_{12} introduces a K-nary potential ~~of~~ connecting $C_{13}, C_{14} \ldots C_{1K}, C_{21}$ & C_{22}.
Thus variable elimination & K-nary potential introduction goes on ~~&~~ till the end.
Thus treewidth $= K$.