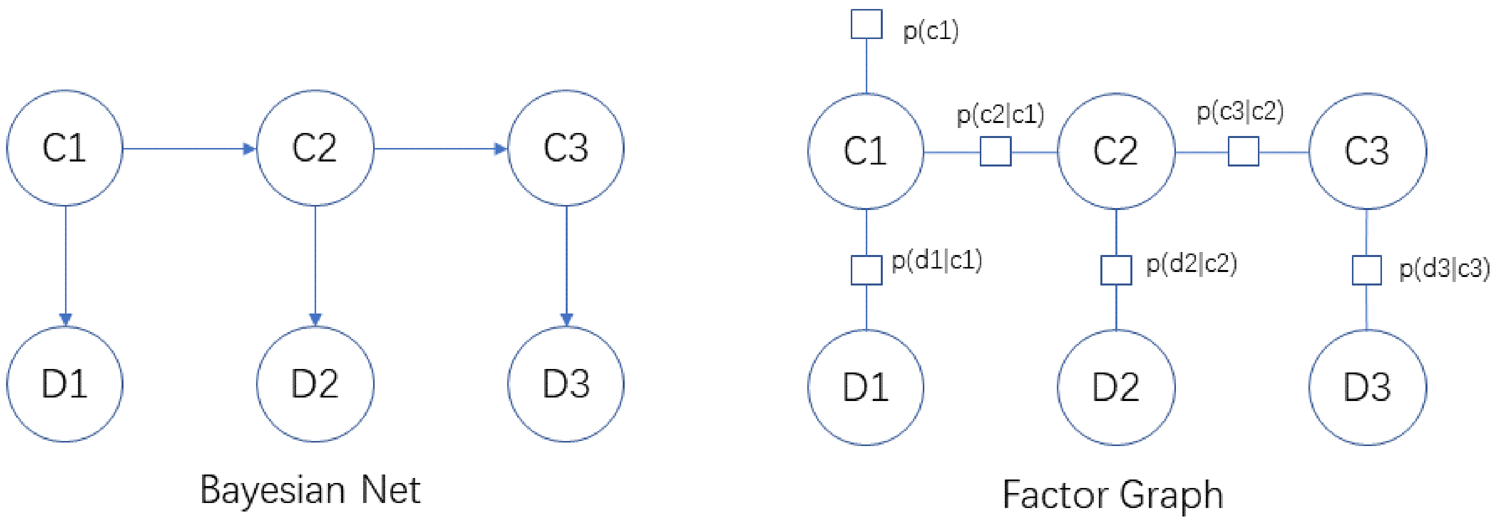
CS221, Spring 2019, PS7 Car

Haiyuan Mei (hmei0411@stanford.edu)

Problem 1: Bayesian network basics

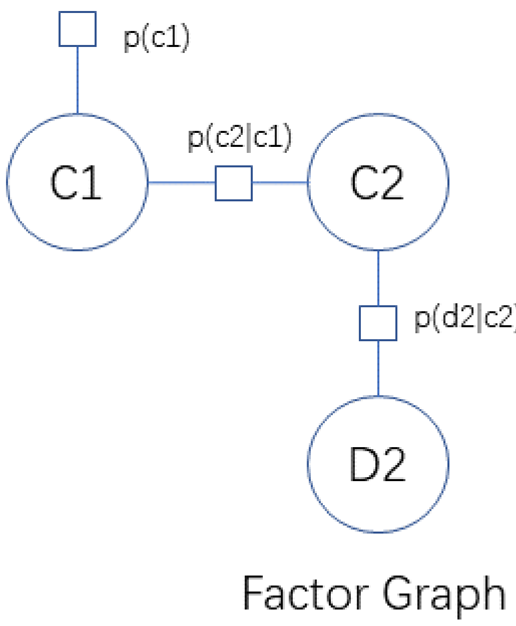
For this problem, the initial Bayesian network and it's factor graph is shown as below:



a. Query $\mathbb{P}(C_2 = 1 | D_2 = 0)$.

Apply the general strategy described in lecture: marginalize non-ancestral variables, condition, and perform variable elimination.

1. After marginalize variables that are not ancestors of Q or E, the factor graph is:



2. Condition on D_2=0 will remove variable D_2, replace the binary factor p(d2|c2) with p(d2=0|c2);
3. Variable elimination. In this case variable C_1 needs to be eliminated, leaving only one variable C_2 and one unary factor f(c2):

$$\begin{aligned} f(c2) &= \sum_{c1} p(c1)p(c2|c1) \\ &= \frac{1}{2}(p(c2|c1=0) + p(c2|c1=1)) \end{aligned}$$

4. The final query $\mathbb{P}(C_2 = 1 | D_2 = 0)$ is hence the product of the factors from 2,3:

$$\begin{aligned} \mathbb{P}(C_2 = 1 | D_2 = 0) &\propto \frac{1}{2}(p(c2=1|c1=0) + p(c2=1|c1=1))p(d2=0|c2=1) \\ &= \frac{1}{2}(\epsilon + (1-\epsilon))\eta \\ &= \frac{1}{2}\eta \\ \mathbb{P}(C_2 = 0 | D_2 = 0) &\propto \frac{1}{2}(p(c2=0|c1=0) + p(c2=0|c1=1))p(d2=0|c2=0) \\ &= \frac{1}{2}(\epsilon + (1-\epsilon))(1-\eta) \\ &= \frac{1}{2}(1-\eta) \end{aligned}$$

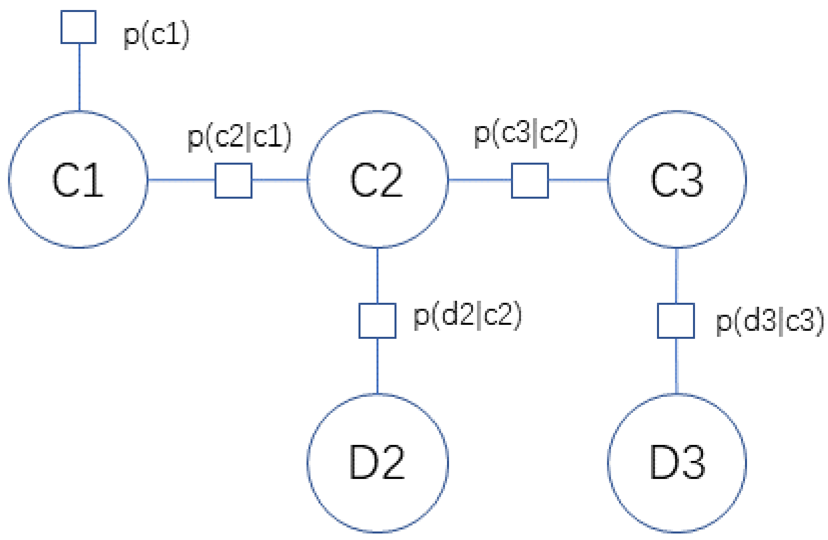
Normalize the probability and get the final result:

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

b. Query $\mathbb{P}(C_2 = 1|D_2 = 0, D_3 = 1)$.

Apply the general strategy described in lecture, marginalize non-ancestral variables, condition, and perform variable elimination.

1. After marginalization, the factor graph looks as following, only C_1 can be removed at this step:



Factor Graph

2. Condition on both $D_2 = 0$ and $D_3 = 1$ will remove variable D_2 and D_3, the corresponding factors get changed to p(d2=0|c2) and p(d3=1|c3)
3. Variable elimination. Both C_1 and C_3 can be eliminated. The case of C_1 is the same as in 1.a:

$$\begin{aligned} f(c2) &= \sum_{c1} p(c1)p(c2|c1) \\ &= \frac{1}{2}(p(c2|c1=0) + p(c2|c1=1)) \end{aligned}$$

   The elimination of C_3 creates a unary factor g(c2):

$$\begin{aligned} g(c2) &= \sum_{c3} p(c3|c2)p(d3=1|c3) \\ &= p(c3=0|c2)p(d3=1|c3=0) + p(c3=1|c2)p(d3=1|c3=1) \end{aligned}$$

4. The final query $\mathbb{P}(C_2 = 1|D_2 = 0, D_3 = 1)$ is the product of the 3 unary factors from step 2 and 3:

$$\begin{aligned} \mathbb{P}(C_2 = 1|D_2 = 0, D_3 = 1) \propto & \frac{1}{2}(p(c2=1|c1=0) + p(c2=1|c1=1)) * \\ & p(c3=0|c2=1)p(d3=1|c3=0) + p(c3=1|c2=1)p(d3=1|c3=1) * \\ & p(d2=0|c2=1) \\ = & \frac{1}{2}\eta(1 - \epsilon - \eta + 2\epsilon\eta) \\ \mathbb{P}(C_2 = 0|D_2 = 0, D_3 = 1) \propto & \frac{1}{2}(p(c2=0|c1=0) + p(c2=0|c1=1)) * \\ & p(c3=0|c2=0)p(d3=1|c3=0) + p(c3=1|c2=0)p(d3=1|c3=1) * \\ & p(d2=0|c2=0) \\ = & \frac{1}{2}(1-\eta)(\epsilon + \eta - 2\epsilon\eta) \end{aligned}$$

   Normalize the above two and get the final result:

$$\mathbb{P}(C_2 = 1|D_2 = 0, D_3 = 1) = \frac{\eta(1 - \epsilon - \eta + 2\epsilon\eta)}{\eta(1 - \epsilon - \eta + 2\epsilon\eta) + (1-\eta)(\epsilon + \eta - 2\epsilon\eta)}$$

c. Suppose ϵ=0.1 and η=0.2

1. The above two queries are:

$$\mathbb{P}(C_2=1|D_2=0)=\eta=0.2$$

$$\begin{aligned}\mathbb{P}(C_2=1|D_2=0,D_3=1)&=\frac{\eta(1-\epsilon-\eta+2\epsilon\eta)}{\eta(1-\epsilon-\eta+2\epsilon\eta)+(1-\eta)(\epsilon+\eta-2\epsilon\eta)}\\&=\frac{0.2(1-0.1-0.2+2*0.1*0.2)}{0.2(1-0.1-0.2+2*0.1*0.2)+(1-0.2)(0.1+0.2-2*0.1*0.2)}\\&=\frac{0.148}{0.148+0.208}\\&\approx 0.4157\end{aligned}$$

2. From the above result, adding the second sencor read $D_3=1$ will reinforce the belief of $C_2=1$ by increasing the probability from 0.2 to 0.4157. The intuition is that since the sensor D_2 indicates the distance of C_2 as 0, the chance of $C_2=1$ is small, $\mathbb{P}(C_2=1|D_2=0)$ is only 0.2; By having a sensor at D_3 giving the distance 1, there will be high probability that C_3 will be 1, hence since C_2 and C_3 should be close to each other, the probability of $\mathbb{P}(C_2=1|D_2=0,D_3=1)$ should be therefore higher.
3. In order to make $\mathbb{P}(C_2=1|D_2=0)=\mathbb{P}(C_2=1|D_2=0,D_3=1)$ solve the equation:

$$\frac{\eta(1-\epsilon-\eta+2\epsilon\eta)}{\eta(1-\epsilon-\eta+2\epsilon\eta)+(1-\eta)(\epsilon+\eta-2\epsilon\eta)}=\eta$$

   Replace η with 0.2:

$$\frac{0.2(1-\epsilon-0.2+2*0.2\epsilon)}{0.2(1-\epsilon-0.2+2*0.2\epsilon)+(1-0.2)(\epsilon+0.2-2*0.2\epsilon)}=0.2$$

   We get $\epsilon=0.5$.

Problem 5: Which car is it?

a. Expression for the conditional distribution $\mathbb{P}(C_{11},C_{12}|E_1=e_1)$

There are two permutations for E_1 as shown below:

For one of the permutation in $e_1=(e_{11},e_{12})$, the conditional distribution is proportional to:

$$p(c_{11})p(c_{12})p_N(e_{11},||a_1-c_{11}||,\delta^2)p_N(e_{12},||a_1-c_{12}||,\delta^2)$$

The two different arrangements of E_1 are basically the same structure, each with the same conditional distribution. So the final conditional distributional should be proportional to the sum of all arrangements.

$$\mathbb{P}(C_{11},C_{12}|E_1=e_1)\propto 2p(c_{11})p(c_{12})p_N(e_{11},||a_1-c_{11}||,\delta^2)p_N(e_{12},||a_1-c_{12}||,\delta^2)$$

b. Show that the number of assignments for all K cars that obtain the maximum value of $\mathbb{P}$ is at least K!.

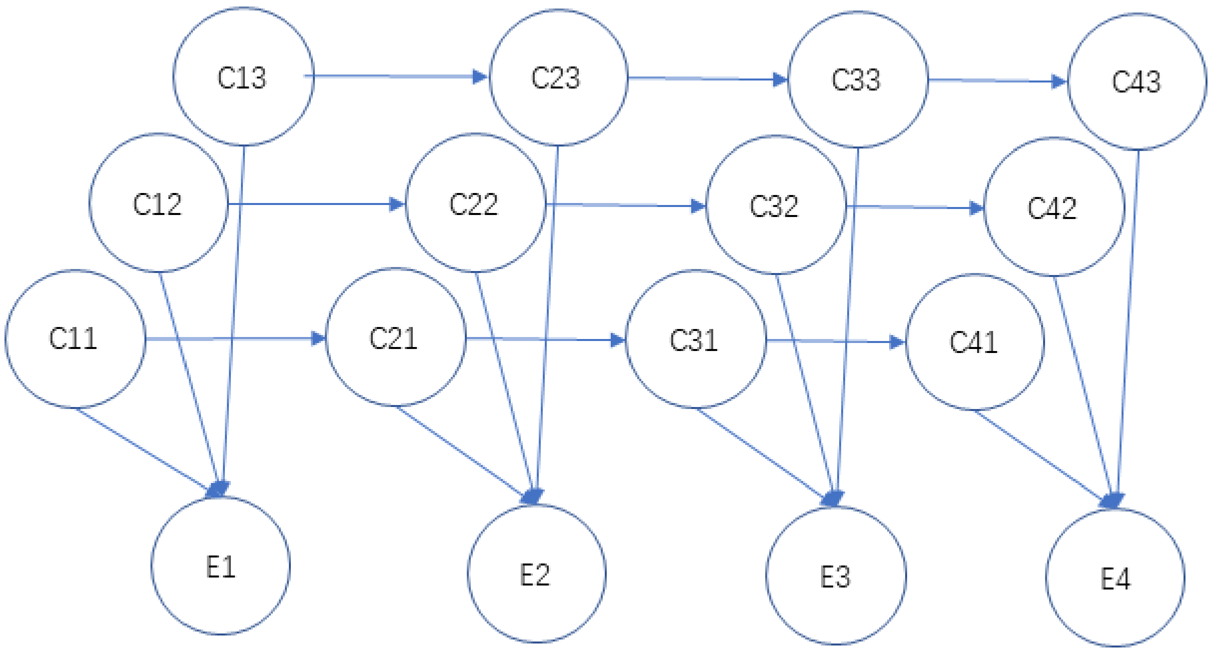
In order to maximize the value of $\mathbb{P}(C_{11},\ldots,C_{1K}|E_1=e_1)$, for each car location C_{1i}, there is a sensor location in E_1, denoted as E_{1j} which maximizes the PDF $p_N(e_{1j},||a_1-c_{1i}||,\delta^2)$

The order of elements in $c_{11},\ldots,c_{1K}$ doesn't change the fact that for eacn of the car location C_{1i}, there is a sensor location in E_1, which gives the maximum value of $\mathbb{P}(C_{11},\ldots,C_{1K}|E_1=e_1)$.

The number of different permutation is K!, number of assignments for all K cars that obtain the maximum value of $\mathbb{P}$ is at least K!.

c. Treewidth corresponding to the posterior distribution over all K car locations

The Bayesian network can be shown as below(for example: K=3,T=4):



If condition on E_i, it will create a K-ary factor for each time step.
Eleminate a car position from left to right, or left to right will both create new k-ary factors.
For this reason, the tree width corresponding to the posterior

$$\mathbb{P}(C_{11} = c_{11}, \dots, C_{1K} = c_{1K}, \dots, C_{T1} = c_{T1}, \dots, C_{TK} = c_{TK} \mid E_1 = e_1, \dots, E_T = e_T)$$

should be K.

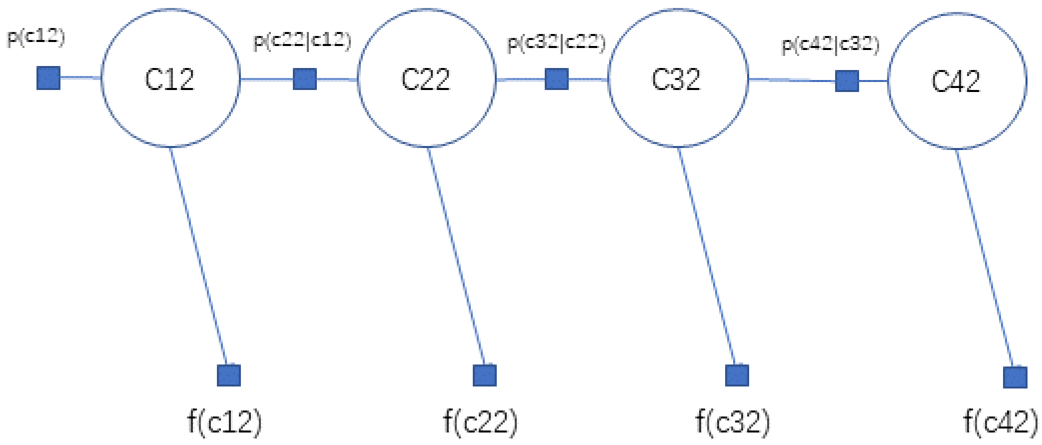
d. Extra

Similar to bayesian network in c. Difference is that element number in E_t is now only K instead of K!.

1. Condition on $E_1 = e_1, \dots, E_T = e_T$, this will creates a new K-ary factor for each of the T times.
2. Each of the K-ary factor in time i can be denoted as:

$$f(e_i|c_{i1}, \dots, c_{ik}) = K \prod_{j=1}^{K} p_N(e_{ij}, ||a_i - c_{ij}||, \delta^2)$$

3. Since the query is only for one car: $p(c_{ti} \mid e_1, \dots, e_T)$, it can be in any index of E_i, there are K different positions. Use this fact the factor graph can be simplified as:



where the f factors are:

$$f(c_{i*}) = \prod_{j=1}^{K} p_N(e_{ij}, ||a_i - c_{ij}||, \delta^2)$$

4. Now it's just a simple HMM model and can use normal variable elimination methods to solve.