CS221 Autumn 2016 Homework: Car

SUNet ID: [06033811]
Name: [Chao Xu]

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

Problem 1: Warm-up

(a) Suppose we have a sensor reading for the second timestep, $D_2 = 0$. Compute the posterior distribution $\mathbf{P}(C_2 = 1|D_2 = 0)$. We encourage you to draw out the (factor) graph.

Follow the general strategy in the class, we can remove all the variables except C_2, and D_2, And the factor graph would looks like the following:

—- o —- C_2 – o – D_2
|
o
|
—- o —- C_1
|
o
|

where on the node left to C_2, it should be local probability: $p(C_2)$, and for the node between C_2 and D_2, it should be $p(D_2|C_2)$, and for the node between C_1 and C_2, it should be $p(C_2|C_1)$. Therefore, by removing C_1 from the factor graph:

$$f(C_2 = 1) = \sum_{C_1} p(C_2 = 1|C_1)p(C_1) = 0.5$$

And for the same procedure, $f(C_2 = 0)$. For the last procedure, we have:

$$\mathbf{P}(C_2 = 1|D_2 = 0) = \frac{p(C_2 = 1)p(D_2 = 0|C_2 = 1)f(C_2 = 1)}{p(C_2 = 1)p(D_2 = 0|C_2 = 1)f(C_2 = 1) + p(C_2 = 1)p(D_2 = 0|C_2 = 0)f(C_2 = 1)}$$

$$\mathbf{P}(C_2 = 1|D_2 = 0) = \frac{0.25\eta}{0.25\eta + 0.25(1-\eta)} = \eta$$

(b) Suppose a time step has elapsed and we got another sensor reading, $D_3 = 1$, but we are still interested in C_2. Compute the posterior distribution $\mathbf{P}(C_2 = 1|D_2 = 0, D_3 = 1)$. The resulting expression might be moderately complex. We encourage you to draw out the (factor) graph.

The factor graph looks like the following by marginalize non-ancestral variables:

```
—- o —- C3 – o – D3
        |
        o
        |
—- o —- C2 – o – D2
        |
        o
        |
        C1
        |
        o
        |
```

And then condition on D_2 and D_3. The node between C_2 and D_2 would be $p(D_2 = 0|C_2)$, and the node between C_3 and D_3 would be $p(D_3 = 1|C_3)$. In addition, the node between C_3 and C_2 should be $p(C_3|C_2)$, and node between C_2 and C_1 should be $p(C_2|C_1)$. Therefore:

$$\mathbf{P}(C_2 = 0|D_2 = 0, D_3 = 1) \propto \sum_{C_1,C_3} p(C_1)p(C_2 = 0)p(C_3)p(C_3|C_2 = 0)p(C_2 = 0|C_1)p(D_3 = 1|C_3)p(D_2 = 0|C_2 = 0)$$
$$\mathbf{P}(C_2 = 0|D_2 = 0, D_3 = 1) \propto 0.25(1 - \eta)(\epsilon + \eta - 2\epsilon\eta)$$

Similarly:

$$\mathbf{P}(C_2 = 1|D_2 = 0, D_3 = 1) \propto \sum_{C_1,C_3} p(C_1)p(C_2 = 1)p(C_3)p(C_3|C_2 = 1)p(C_2 = 1|C_1)p(D_3 = 1|C_3)p(D_2 = 0|C_2 = 1)$$
$$\mathbf{P}(C_2 = 1|D_2 = 0, D_3 = 1) \propto 0.25\eta(1 - \epsilon - \eta + 2\epsilon\eta)$$

After normalization:

$$\mathbf{P}(C_2 = 1|D_2 = 0, D_3 = 1) = \frac{0.25\eta(1-\epsilon-\eta+2\epsilon\eta)}{0.25\eta(1-\epsilon-\eta+2\epsilon\eta)+0.25(1-\eta)(\epsilon+\eta-2\epsilon\eta)}$$

(c) Suppose ϵ=0.1 and η=0.2.

i. Compute and compare the probabilities $\mathbf{P}(C_2 = 0|D_2 = 1)$ and $\mathbf{P}(C_2 = 1|D_2 = 0, D_3 = 1)$. Give numbers, round your answer to 4 significant digits.

$\mathbf{P}(C_2 = 1|D_2 = 0, D_3 = 1) = \dfrac{0.25\eta(1-\epsilon-\eta+2\epsilon\eta)}{0.25\eta(1-\epsilon-\eta+2\epsilon\eta)+0.25(1-\eta)(\epsilon+\eta-2\epsilon\eta)}$

$\mathbf{P}(C_2 = 1|D_2 = 0, D_3 = 1) = \dfrac{0.148}{0.356} = 0.4157$

For $\mathbf{P}(C_2 = 0|D_2 = 1)$:

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

ii. How did adding the second sensor reading $D_3 = 1$ change the result? Explain your intuition in terms of the car positions with respect to the observations.

Adding the second sensor reading $D_3 = 1$ double the probability of $C_2 = 1$. The reason is that adding the D_3 observation enhanced the probability of $C_3 = 1$, which enhanced the probability of C_2.

iii. What would you have to set ϵ while keeping η=0.2 so that $\mathbf{P}(C_2 = 1D_2 = 0) = \mathbf{P}(C_2 = 1D_2 = 0, D_3 = 1)$? Explain your intuition in terms of the car positions with respect to the observations.

ϵ should equal to 0.5. Because the car has same probability for moving or not moving, and therefore observing next position tells us nothing about the previous position.

Problem 5: Which car is it?

(a) Suppose we have K=2 cars and one time step T=1. Write an expression for the conditional distribution $\mathbf{P}(C_{11}, C_{12}|E_1 = e_1)$ as a function of the PDF of a Gaussian $p(v; \mu, \sigma_2)$ and the prior probability $p(c_{11})$ and $p(c_{12})$ over car locations. Your final answer should not contain variables d_{11}, d_{12}.

To determine the conditional distribution, the Gaussian should be computed for each permutation and summed over the evidence. The result should be as following:

$p(c_{11}, c_{12}|e_1) \propto (2p(c_{11})p_{\mathcal{N}}(e_{11}; \|a_1 - c_{11}\|; \sigma^2)p(c_{12})p_{\mathcal{N}}(e_{12}; \|a_1 - c_{12}\|; \sigma^2))$

(b) Assuming the prior $p(c_{1i})$ is the same for all i, show that the number of assignments for all K cars $(c_{11}, , c1K)$ that obtain the maximum value of $\mathbf{P}(C_{11} = c_{11}, , C_{1K} = c_{1K}|E_1 = e_1)$ is at least $K!$.

Car locations that maximize the probability are unique ($C_{1K} \neq c_{1K}$ for all i $\neq$ j) and assume there is one assignment of C_{1K} that reaches the maximum value. Through permutation, there exists $K!$ assignments of C_{1K} that can reach the same maximum value. Therefore, the number of assignments for all K cars $(c_{11}, , c1K)$ that obtain the maximum value of $\mathbf{P}(C_{11} = c_{11}, , C_{1K} = c_{1K} | E_1 = e_1)$ is at least $K!$.

(c) For general K, what is the treewidth corresponding to the posterior distribution over all K car locations at all T time steps conditioned on all the sensor readings.

By conditioning on all the sensor readings and using general strategy in the lecture, the remaining number of variables in the scope would be K cars at time T. Therefore, the treewidth should be K.