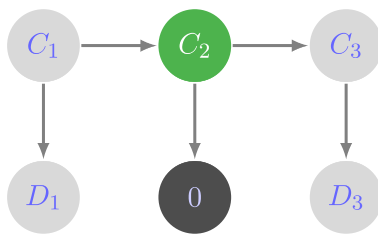
Homework 7: Car Tracking

Brandon McKinzie

Problem 1: Bayesian Network Basics

(a) *Suppose we have a sensor reading for the second timestep, $D_2 = 0$. Compute the posterior distribution $\mathbb{P}(C_2 = 1 \mid D_2 = 0)$.*

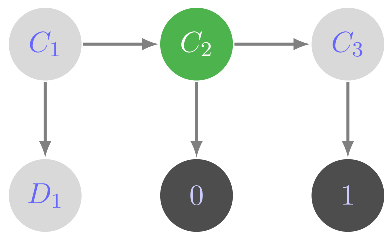
Below is the Bayesian network, where we've observed $D_2 = 0$:



$$
\begin{aligned}
\Pr\left[C_2 = 1 \mid D_2 = 0\right] &\propto \Pr\left[C_2 = 1, D_2 = 0\right] && (1)\\
&= \sum_{c_1} \Pr\left[C_2 = 1, D_2 = 0, c_1\right] && (2)\\
&= \sum_{c_1} \Pr\left[c_1\right] \Pr\left[C_2 = 1 \mid c_1\right] \Pr\left[D_2 = 0 \mid C_2 = 1\right] && (3)\\
&= 0.5 \sum_{c_1} \Pr\left[C_2 = 1 \mid c_1\right] \Pr\left[D_2 = 0 \mid C_2 = 1\right] && (4)\\
&= 0.5\eta \sum_{c_1} \Pr\left[C_2 = 1 \mid c_1\right] && (5)\\
&= 0.5\eta(\epsilon + (1-\epsilon)) && (6)\\
&= 0.5\eta && (7)\\
\Pr\left[D_2 = 0\right] &= \sum_{c_2} \Pr\left[D_2 = 0, c_2\right] && (8)\\
&= \sum_{c_2} \Pr\left[D_2 = 0 \mid c_2\right] \sum_{c_1} \Pr\left[c_2 \mid c_1\right] \Pr\left[c_1\right] && (9)\\
&= (1-\eta) \cdot (0.5(1-\epsilon) + 0.5\epsilon) + \eta \cdot (0.5\epsilon + 0.5(1-\epsilon)) && (10)\\
&= 0.5 && (11)\\
\therefore \Pr\left[C_2 = 1 \mid D_2 = 0\right] &= \frac{\Pr\left[C_2 = 1, D_2 = 0\right]}{\Pr\left[D_2 = 0\right]} = \eta && (12)
\end{aligned}
$$

(b) *Compute* $\mathbb{P}(C_2 = 1 \mid D_2 = 0, D_3 = 1)$

Now our Bayesian network looks like:



$$
\begin{aligned}
p(C_2, D_2, D_3) &= \sum_{c_1} \sum_{c_3} p(c_1) p(c_2 \mid c_1) p(d_2 \mid c_2) p(c_3 \mid c_2) p(d_3 \mid c_3) && (13)\\
&= 0.5 \sum_{c_1} \sum_{c_3} p(c_2 \mid c_1) p(d_2 \mid c_2) p(c_3 \mid c_2) p(d_3 \mid c_3) && (14)\\
&= 0.5 \cdot p(d_2 \mid c_2) \sum_{c_3} p(c_3 \mid c_2) p(d_3 \mid c_3) && (15)\\
p(C_2 = 1, D_2 = 0, D_3 = 1) &= 0.5\eta \left(\epsilon\eta + (1-\epsilon)(1-\eta)\right) && (16)\\
p(D_2 = 0, D_3 = 1) &= \sum_{c_2} p(c_2, D_2 = 0, D_3 = 1) && (17)\\
&= 0.5\eta \left(\epsilon\eta + (1-\epsilon)(1-\eta)\right) + 0.5(1-\eta)\left((1-\epsilon)\eta + \epsilon(1-\eta)\right) && (18)\\
p(C_2 = 1 \mid D_2 = 0, D_3 = 1) &= \frac{0.5\eta \left(\epsilon\eta + (1-\epsilon)(1-\eta)\right)}{0.5\eta \left(\epsilon\eta + (1-\epsilon)(1-\eta)\right) + 0.5(1-\eta)\left((1-\epsilon)\eta + \epsilon(1-\eta)\right)} && (19)\\
&= \frac{\epsilon\eta^2 + \eta(1-\epsilon)(1-\eta)}{\epsilon\eta^2 + +2(1-\eta)(1-\epsilon)\eta + \epsilon(1-\eta)^2} && (20)
\end{aligned}
$$

(c)

i.

$$P(C_2 = 1 \mid D_2 = 0) = \eta = 0.2 \tag{21}$$

$$P(C_2 = 1 \mid D_2 = 0, D_3 = 1) = \frac{(0.1)(0.2)^2 + 0.2(0.9)(0.8)}{(0.1)(0.2)^2 + 2(0.8)(0.9)(0.2) + (0.1)(0.8)^2} \tag{22}$$

$$\approx 0.4157 \tag{23}$$

ii. The second reading ($D_3 = 1$) makes it more likely that $C_2 = 1$. Informally, after we observe $D_3 = 1$ while realizing that it's unlikely the car has actually moved (relative to the probability that our sensor was wrong), it becomes more probable that our previous reading of $D_2 = 0$ was just a bad reading.

iii. We can compute the value of ϵ by equating the two formulas and solving:

$$\eta = \frac{\epsilon\eta^2 + \eta(1-\epsilon)(1-\eta)}{\epsilon\eta^2 + +2(1-\eta)(1-\epsilon)\eta + \epsilon(1-\eta)^2} \tag{24}$$

$$1 = \frac{\epsilon\eta + (1-\epsilon)(1-\eta)}{\epsilon\eta^2 + +2(1-\eta)(1-\epsilon)\eta + \epsilon(1-\eta)^2} \tag{25}$$

$$\epsilon\eta + (1-\epsilon)(1-\eta) = \epsilon\eta^2 + +2(1-\eta)(1-\epsilon)\eta + \epsilon(1-\eta)^2 \tag{26}$$

$$\epsilon\eta = \epsilon\eta^2 + +(2\eta - 1)(1-\eta)(1-\epsilon) + \epsilon(1-\eta)^2 \tag{27}$$

$$\epsilon(\eta - \eta^2 - (1-\eta)^2) = (2\eta - 1)(1-\eta)(1-\epsilon) \tag{28}$$

$$\epsilon((1-\eta)(\eta - (1-\eta))) = (2\eta - 1)(1-\eta)(1-\epsilon) \tag{29}$$

$$\epsilon(1-\eta)(2\eta - 1) = (2\eta - 1)(1-\eta)(1-\epsilon) \tag{30}$$

$$\epsilon = 1 - \epsilon \tag{31}$$

$$\epsilon = \frac{1}{2} \tag{32}$$

Intuitively, we'd have to set $\epsilon = \frac{1}{2}$ (car equally likely to be in 0 or 1 independent of previous location) since that would make additional observations essentially useless.

Problem 5: Which car is it?

- *$C_{ti} \in \mathbb{R}^2$: location of ith car at time t, where $1 \leq i \leq K$, and $1 \leq t \leq T$.*
- *$D_t = \{D_{t1}, \ldots, D_{tK}\}$, where each $D_{ti} \in \mathbb{R}$ is noisy distance measurement of ith car at time t.*

(a) *Write an expression for Pr $[C_{11}, C_{12} \mid E_1 = e_1]$ as a function of $\mathcal{N}(v; \mu, \sigma^2)$ and the priors $p(c_{11})$ and $p(c_{12})$.*

Since $T = 1$, I'm going to drop the time index in the following calculations to avoid confusing myself. Then, since $K = 2$, I'll denote $E_1 \equiv E \equiv (E_1, E_2)$ (again, I've dropped the time index since it is always 1 for this problem). So we need an expression for $\Pr[C_1, C_1 \mid E_1, E_2]$.

First, we can rewrite

$$\Pr[C_1, C_1 \mid E_1, E_2] \propto \Pr[E_1, E_2 \mid C_1, C_2] \Pr[C_1, C_2] \tag{33}$$
$$= \Pr[E_1, E_2 \mid C_1, C_2] \Pr[C_1] \Pr[C_2] \tag{34}$$

where we've taken advantage of the independence of the two cars. Now we need to break down $\Pr[E_1, E_2 \mid C_1, C_2]$. Since E is sampled uniformly at random from the possible permutations in D, we can think of a given instantiation $E = (e_1, e_2)$ as an event whose probability is proportional to the union of all possible readings (d_1, d_2) that could result in (e_1, e_2). Formally,

$$\Pr[E_1 = e_1, E_2 = e_2 \mid C_1, C_2] \propto \Pr[D_1 = e_1, D_2 = e_2 \mid C_1, C_2] + \Pr[D_1 = e_2, D_2 = e_1 \mid C_1, C_2] \tag{35}$$
$$= \Pr[D_1 = e_1 \mid C_1] \Pr[D_2 = e_2 \mid C_2] + \Pr[D_1 = e_2 \mid C_1] \Pr[D_2 = e_1 \mid C_2] \tag{36}$$

where again I've used the independence of the two cars in the final line. We can replace the $D_i \mid C_i$ conditionals with the provided normal distribution to get the final expression (with time index included). Let $\mathcal{N}_{C_{ti}}(x) = \mathcal{N}(x; ||a_t - C_{ti}||; \sigma^2)$ (for brevity's sake).

$$\Pr[C_{11}, C_{12} \mid E_1 = e_1] \propto \Pr[C_{11}] \Pr[C_{12}] \left[\mathcal{N}_{C_{11}}(e_{11})\mathcal{N}_{C_{12}}(e_{12}) + \mathcal{N}_{C_{11}}(e_{12})\mathcal{N}_{C_{12}}(e_{11})\right] \tag{37}$$

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

Intuitively, by introducing the constraint that $p(c_{1i})$ is the same for all i, combined with the fact that we don't know which car each element of E_1 corresponds to, we can take any of the assignments that maximize $\Pr[C_{11}, \ldots, C_{1K} \mid E_1]$ and permute the locations of the cars without changing the probability. Since there are $K!$ such permutations, there are at least $K!$ assignments of the cars that obtain the maximum value.

(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:*

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

Given the factor graph, we'll first have to condition on the evidence $e_1, \ldots, e_T$. This will create factors f_t for all timesteps $1 \leq t \leq T$, each with arity K. Since we want a variable ordering that starts with variables with the least amount of neighbors, we'll order along timesteps (from 1 to T). Since the Markov blanket of each variable we eliminate along timesteps has K variables (corresponding to the $K - 1$ other cars plus the transition), the resultant factors will have arities no larger than K. By definition, then, the treewidth is K.

(d) *Now suppose you change your sensors so that at each time step t, they return the list of exact positions of the K cars, but shifted (with wrap around) by a random amount. For example, if the true car positions at time step 1 are* $c_{11} = 1, c_{12} = 3, c_{13} = 8, c_{14} = 5$*, then* e_1 *would be* $[1, 3, 8, 5], [3, 8, 5, 1], [8, 5, 1, 3]$*, or* $[5, 1, 3, 8$*, each with probability 1/4. Describe an efficient algorithm for computing* $p(c_{ti} \mid e_1, \ldots, e_T)$ *for any time step t and car i. Your algorithm should not be exponential in K or T.*