CS 440 - Assignment 3

Cameron Pascal, Ariel Shaposhnik

October 18, 2017

Question 1

a

$$P(a, b, c, d, e) = P(a)P(b)P(c)P(d \mid a, b)P(e \mid b, c)$$
$$= 0.2 \cdot 0.5 \cdot 0.8 \cdot 0.1 \cdot 0.3$$
$$= 0.0024$$

b

$$P(\neg a, \neg b, \neg c, \neg d, \neg e) = P(\neg a)P(\neg b)P(\neg c)P(\neg d \mid \neg a, \neg b)P(\neg e \mid \neg b, \neg c)$$
$$= 0.8 \cdot 0.5 \cdot 0.2 \cdot 0.1 \cdot 0.8$$
$$= 0.0064$$

c

$$\mathbf{P}(A \mid b, c, d, e) = \alpha \, \mathbf{P}(A, b, c, d, e)$$

$$P(\neg a, b, c, d, e) = P(\neg a)P(b)P(c)P(d \mid \neg a, b)P(e \mid b, c)$$
$$= 0.8 \cdot 0.5 \cdot 0.8 \cdot 0.6 \cdot 0.3$$
$$= 0.0576$$

$$\mathbf{P}(A \mid b, c, d, e) = \alpha \, \langle 0.0024, 0.0576 \rangle$$

$$\alpha = \frac{1}{0.0024 + 0.0576} = \frac{1}{0.06}$$

$$P(\neg a \mid b, c, d, e) = \frac{0.0576}{0.06} = 0.96$$

Question 2

a

$$\mathbf{P}(B \mid j, m) = \alpha \, \mathbf{P}(B) \sum_e P(e) \, \mathbf{P}(a \mid B, e) P(j \mid a) P(m \mid a)$$

$$\mathbf{P}(B) = \mathbf{f}_1(B) \quad \mathbf{P}(E) = \mathbf{f}_2(E) \quad \mathbf{P}(A \mid B, E) = \mathbf{f}_3(A, B, E)$$

$$\mathbf{P}(j \mid A) = \mathbf{f}_4(A) \quad \mathbf{P}(m \mid A) = \mathbf{f}_5(A)$$

$$\begin{aligned}
\mathbf{f}_6(B, E) &= \sum_a \mathbf{f}_3(A, B, E) \times \mathbf{f}_4(A) \times \mathbf{f}_5(A) \qquad " \times " \text{ is the pointwise product operator} \\
&= [\mathbf{f}_3(a, B, E) \times \mathbf{f}_4(a) \times \mathbf{f}_5(a)] + [\mathbf{f}_3(\neg a, B, E) \times \mathbf{f}_4(\neg a) \times \mathbf{f}_5(\neg a)] \\
&= \left(0.90 \times 0.70 \times \begin{bmatrix} 0.95 & 0.94 \\ 0.29 & 0.001 \end{bmatrix}\right) + \left(0.05 \times 0.01 \times \begin{bmatrix} 0.05 & 0.06 \\ 0.71 & 0.999 \end{bmatrix}\right) \\
&= \begin{bmatrix} 0.5985 & 0.5922 \\ 0.1827 & 6.3 \cdot 10^{-4} \end{bmatrix} + \begin{bmatrix} 2.5 \cdot 10^{-5} & 3 \cdot 10^{-5} \\ 3.55 \cdot 10^{-4} & 4.995 \cdot 10^{-4} \end{bmatrix} \\
&= \begin{bmatrix} 0.598525 & 0.59223 \\ 0.183055 & 0.0011295 \end{bmatrix}
\end{aligned}$$

$$\mathbf{P}(B \mid j, m) = \alpha \, \mathbf{f}_1(B) \times \sum_e \mathbf{f}_2(e) \times \mathbf{f}_6(B, e)$$

$$\begin{aligned}
\mathbf{f}_7(B) &= \sum_e \mathbf{f}_2(e) \times \mathbf{f}_6(B, e) \\
&= [\mathbf{f}_2(e) \times \mathbf{f}_6(B, e)] + [\mathbf{f}_2(\neg e) \times \mathbf{f}_6(B, \neg e)] \\
&= \left(0.002 \times \begin{bmatrix} 0.598525 \\ 0.183055 \end{bmatrix}\right) + \left(0.998 \times \begin{bmatrix} 0.59223 \\ 0.0011295 \end{bmatrix}\right) \\
&= \begin{bmatrix} 0.5922300318 \\ 0.001493351 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(B \mid j, m) &= \alpha \, \mathbf{f}_1(B) \times \mathbf{f}_7(B) \\
&= \alpha \begin{bmatrix} 0.001 \cdot 0.5922300318 \\ 0.999 \cdot 0.001493351 \end{bmatrix} \\
&= \alpha \begin{bmatrix} 5.922300318 \cdot 10^{-4} \\ 0.0014918576 \end{bmatrix}
\end{aligned}$$

$$\alpha = \frac{1}{5.922300318 \cdot 10^{-4} + 0.0014918576} = 0.0020840876^{-1}$$

$$\mathbf{P}(B \mid j, m) = \begin{bmatrix} 0.02841676703 \\ 0.7165490548 \end{bmatrix}$$

b For exact inference by variable elimination there are: 16 multiplications and 6 additions to find the distribution. There is 1 addition and 2 divisions to normalize the distribution, making the total arithmetic operations to be $16 + 6 + 1 + 2 = 25$.

For exact inference by enumeration there are: 15 multiplications and 3 additions for each value of B totaling $2(15 + 3) = 36$ arithmetic operations to find the distribution. There is 1 addition and 2 divisions to normalize the distribution, increasing the total to $36 + 1 + 2 = 39$ arithmetic operations.

c It is known that exact inference by enumeration in Bayesian networks is always $O(2^n)$ where n is the number of random variables represented by the Bayesian network. However, using variable elimination to perform exact inference on a *chain*, the time complexity can be improved to run in linear time. Notice that a *chain* is simply a *polytree* (i.e. there is exactly one path between any two vertices in the network). It is known that variable elimination of polytrees is linear with respect to the number of CPT entries in the network. Given a *chain* with n random variables there is one CPT with 2 entries (corresponding to the root) and another $n-1$ CPTs with 4 entries each (each random variable has only one parent). This gives the total number of CPT entries to be $4(n-1) + 2 = 4n - 4 + 2 = 4n - 2$ and it follows that the complexity of variable elimination on a *chain* is $O(n)$.

Question 3

a Let $Y = Y_1, \cdots, Y_n$ be the set children of the variable X and let any Z_{ij} be the *j-th* parent of Y_i

$$\begin{aligned}
\mathbf{P}(X_i \mid \mathrm{MB}(X_i)) &= \mathbf{P}(X_i \mid \mathrm{parents}(X_i), Y, Z_{1j}, \cdots, Z_{jn}) \\
&= \alpha\, \mathbf{P}(X_i \mid \mathrm{parents}(X_i), Z_{1j}, \cdots, Z_{jn})\mathbf{P}(Y \mid X_i\ \mathrm{parents}(X_i), Z_{1j}, \cdots, Z_{jn}) \\
&= \alpha\, \mathbf{P}(X_i \mid \mathrm{parents}(X_i))\mathbf{P}(Y \mid X_i, Z_{1j}, \cdots, Z_{jn}) \\
&= \alpha\, \mathbf{P}(X_i \mid \mathrm{parents}(X_i)) \prod_{Y_j \in Y} \mathbf{P}(Y_j \mid \mathrm{parents}(X_i))
\end{aligned}$$

This derivation is possible because in a Bayesian network, a variable X_i is independent of all variables not descending from X_i.

b To find the total number of states that a Gibb's sampling approach (a version of MCMC), consider the evidence variables: *Sprinkler* and *WetGrass* in the query, as well as the non-evidence variables: *Rain* and *Cloudy*. The starting state assignments will be the values of the evidence variables and random values for the non-evidence variables. Because the variables in the network are boolean and there are only 2 non-evidence variables, the number of initial states is the number of value combinations *Sprinkler* and *WetGrass* can be assigned, which is $2^2 = 4$. After initializing the values of the variables in the network, the 2 non-evidence variables to be sampled can only be assigned 2 values each; increasing the total number of states to be $2^2 \cdot 2^2 = 2^4$. After sampling an arbitrary non-evidence variable, there is only 1 other non-evidence variable to sample, which again has only 2 possible assignments. This final assignment makes the total size of the state space to be $2^2 \cdot 2^2 \cdot 2 = 2^5 = 32$. Note, while the number of paths to each state is 2^5, there are only $1 \cdot 1 \cdot 2 \cdot 2 = 4$

distinct assignments to the variables in the state space, as the two evidence variables are fixed, and only the two non-evidence variables are allowed to vary.

c Q will be a 4×4 matrix (there are 4 variables in the network, but two of them are fixed, so only the other two: *Rain* and *Cloudy* will vary, hence there are only 16 possible transitions). For simplicity the variables: *Cloudy*, *Rain*, *WetGrass*, and *Sprinkler* are abbreviated C, R, W, and S. The distributions of the non-evidence variables given their Markov Blankets of will be computed to find the transition matrix.

$$\begin{aligned}
\mathbf{P}(R \mid c, w, s) &= \mathbf{P}(R \mid c)\mathbf{P}(w \mid s, R) \\
&= \alpha \langle 0.792, 0.18 \rangle \\
&= \langle 22/27, 5/27 \rangle
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(R \mid \neg c, w, s) &= \mathbf{P}(R \mid \neg c)\mathbf{P}(w \mid s, R) \\
&= \alpha \langle 0.198, 0.720 \rangle \\
&= \langle 11/51, 40/51 \rangle
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(C \mid r, s) &= \mathbf{P}(C)\mathbf{P}(s \mid C)\mathbf{P}(r \mid C) \\
&= \alpha \langle 0.04, 0.05 \rangle \\
&= \langle 4/9, 5/9 \rangle
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(C \mid \neg r, s) &= \mathbf{P}(C)\mathbf{P}(s \mid C)\mathbf{P}(\neg r \mid C) \\
&= \alpha \langle 0.01, 0.20 \rangle \\
&= \langle 1/21, 20/21 \rangle
\end{aligned}$$

The probability of a transition $P(y \to y')$ is given by the Gibb's sampling algorithm, which is $\mathbf{P}(Z_i \mid mb(Z_i))$ conditioned on the current values of the variables in state y. There are three transition states that must be considered:

(a) There are two variables varying between the two states. The probability of this is 0, as MCMC only allows one change between states.

(b) The transitioned state is the same as the prior state (i.e. $y = y'$). The probability of a transition of this type, is the probability of sampling either non-evidence variable.

$$P((c, r) \to (c, r)) = P(c)[P(r \mid c, w, s) + P(c \mid r, s)]$$

(c) Only one non-evidence variable varies.

$$P((c, r) \to (c, \neg r)) = P(c)P(\neg r \mid c, w, s)$$

Following from these cases the transition matrix Q is constructed as follows: (an entry i, j in Q corresponds to the transition $y_i \to y_j$:

$$Q = \begin{bmatrix} 17/27 & 5/54 & 5/18 & 0 \\ 11/27 & 22/189 & 0 & 10/21 \\ 2/9 & 0 & 59/123 & 20/51 \\ 0 & 1/42 & 11/102 & 310/357 \end{bmatrix} \qquad \begin{aligned} y_1 &= (c, r, w, s) \\ y_2 &= (c, \neg r, w, s) \\ y_3 &= (\neg c, r, w, s) \\ y_4 &= (\neg c, \neg r, w, s) \end{aligned}$$

Question 4

a The net gain of purchasing a car in good quality $q+$ is $\$4000 - \$3000 = \$1000$, similarly for $q-$: $\$4000 - \$3000 - \$1400 = -\$400$. Using $P(q+) = 0.7$ and $P(q-) = 1 - P(q+) = 0.3$, the expected net gain will be:

$$(P(q+) \cdot \$1000) + (P(q-) \cdot -\$400) = \$580$$

b It is given that $P(\text{pass}\,|\,q+) = 0.8$ and $P(\text{pass}\,|\,q-) = 0.35$. To find the likelihood of the car car's quality ($q+$ and $q-$), given the results of the mechanic's inspection, the probability of the car passing inspection as well as the probability of a car failing inspection must be computed.

$$\begin{aligned} P(\text{pass}) &= P(\text{pass}\,|\,q+)P(q+) + P(\text{pass}\,|\,q+)P(q-) \\ &= 0.8 \cdot 0.7 + 0.35 \cdot 0.3 \\ &= 0.665 \\ P(\text{fail}) &= 1 - P(\text{pass}) \\ &= 1 - 0.665 \\ &= 0.335 \end{aligned}$$

With those probabilities known, the posterior probabilities of the car's quality given it's inspection can be found. (The computations are not show, as they are trivial.)

$$\begin{aligned} P(q+\,|\,\text{pass}) &= 0.8421 \\ P(q+\,|\,\text{fail}) &= 0.4179 \\ P(q\text{-}\,|\,\text{pass}) &= 0.1578 \\ P(q\text{-}\,|\,\text{fail}) &= 0.5821 \end{aligned}$$

c To find the expected utility, the expected net gain of purchasing the car and the posterior probabilities calculated earlier are required.

  (a) Expected utility, if the car passes inspection

  $$[P(q+\,|\,\text{pass} \cdot 1000] + [P(q\text{-}\,|\,\text{pass}) \cdot -400] = 778.94$$

  (b) Expected utility, if the car fails inspection

  $$[P(q+\,|\,\text{fail}) \cdot 1000] + [P(q\text{-}\,|\,\text{fail}) \cdot -400] = 185.06$$

In either case, the expected utility is positive and the car should be purchased.

d Having the car inspected is a waste of money and is not valuable. This is because, as explained above, the car should be purchased regardless of whether or not it passes the mechanic's inspection.