Homework 8
Bayes and Markov

Due: 6:00pm on 11/18/09

Problem 8.1

Read Chapter 14 in Russel and Norvig. Pay careful attention to section 14.4.

Problem 8.2

a. For the Bayesian network given in Russel and Norvig, p. 494, apply variable elimination to the query $P(Burglary|JohnCalls = true, MaryCalls = true)$. Follow the walkthrough of calculations on p. 507 and give the result of each intermediate calculation. In other words, provide your computed values for $f_M(A)$, $f_{\bar{A}JM}(B,E)$,$f_{\bar{E}\bar{A}JM}(B)$, and $P(B|j,m)$.

Solution: $\mathbf{f}_M(A) = \begin{pmatrix} P(m \mid a) \\ P(m \mid \neg a) \end{pmatrix} = \boxed{\begin{pmatrix} .70 \\ .01 \end{pmatrix}}$

$$\mathbf{f}_{\bar{A}JM}(B,E) = \mathbf{f}_A(a,B,E) \times \mathbf{f}_J(a) \times \mathbf{f}_M(a) + \mathbf{f}_A(\neg a,B,E) \times \mathbf{f}_J(\neg a) \times \mathbf{f}_M(\neg a)$$
$$= \begin{pmatrix} .95 & .94 \\ .29 & .001 \end{pmatrix} \times (.90) \times (.70) + \begin{pmatrix} .05 & .06 \\ .71 & .999 \end{pmatrix} \times (.05) \times (.01)$$
$$= \begin{pmatrix} .5985 & .5922 \\ .1827 & .00063 \end{pmatrix} + \begin{pmatrix} .000025 & .00003 \\ .000355 & .0004995 \end{pmatrix}$$
$$= \boxed{\begin{pmatrix} .598525 & .59223 \\ .183055 & .0011295 \end{pmatrix}}$$

$$\mathbf{f}_{\bar{E}\bar{A}JM}(B) = \mathbf{f}_E(e) \times \mathbf{f}_{\bar{A}JM}(B,e) + \mathbf{f}_E(\neg e) \times \mathbf{f}_{\bar{A}JM}(B,\neg e)$$
$$=(.002) \times \begin{pmatrix} .598525 \\ .183055 \end{pmatrix} + (.998) \times \begin{pmatrix} .59223 \\ .0011295 \end{pmatrix}$$
$$= \begin{pmatrix} .001197 \\ .000366 \end{pmatrix} + \begin{pmatrix} .591 \\ .00113 \end{pmatrix}$$
$$= \boxed{\begin{pmatrix} .592 \\ .0015 \end{pmatrix}}$$

$$\mathbf{P}(B \mid j,m) = \alpha \mathbf{f}_B(B) \times \mathbf{f}_{\bar{E}\bar{A}JM}(B) = \mathbf{f}_B(B) \times \alpha \mathbf{f}_{\bar{E}\bar{A}JM}(B)$$
$$= \alpha \begin{pmatrix} .001 \\ .999 \end{pmatrix} \times \begin{pmatrix} .592 \\ .0015 \end{pmatrix} = \alpha \begin{pmatrix} .000592 \\ .0015 \end{pmatrix}$$

$$= \boxed{\begin{pmatrix} .283 \\ .717 \end{pmatrix}}$$

b. Prove that the complexity of a running variable elimination on a polytree network is linear in the size of the tree for any variable ordering consistent with the network structure.

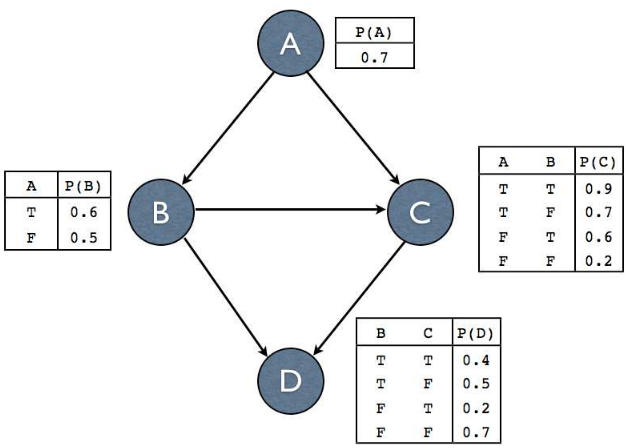
Solution:

Base Case: If there is only 1 element in the polytree network, it will take linear time because there is only one operation to be computed in the polytree network.

Inductive Step: Let n be a polytree network which takes linear time to run variable elimination. Adding another element to the network will add either an addition node or a multiplication node. Both of these nodes will add a constant factor to the computation since the rely only on the node(s) under them to compute a value. Thus adding a node to a network which can be computed in linear time will create a new network which can also be computed in linear time.

Since a network of size k can be computed in linear time if there is a network of size *k-1* which can be computed in linear time, and a network of size 1 can be computed in linear time, every network can be computed in linear time.

Problem 8.3



For the network given above:

a. Write out the formula to compute $P(A = T|D = F)$.

b. Compute the value of $P(A = T|D = F)$.

Solution:

a.

$$\begin{aligned}P(A|\bar{D}) &= \frac{P(A \wedge \bar{D})}{P(\bar{D})} \\ &= \frac{P(A \wedge \bar{D})}{P(A \wedge \bar{D}) + P(\bar{A} \wedge \bar{D})} \\ &= \frac{P(A)P(\bar{D}|A)}{P(A)P(\bar{D}|A) + P(\bar{A})P(\bar{D}|\bar{A})} \\ &= \frac{P(A)\sum_b P(b|A)\sum_c P(c|A,b)P(D|b,c)}{\sum_a P(a)\sum_b P(b|a)\sum_c P(c|a,b)P(\bar{D}|b,c)}\end{aligned}$$

b. 0.749041

Problem 8.4

Consider a Markov model with the state space $S = \{1A, 1B, 2A, 2B, 3A, 3B, T\}$ and action set $A = \{\alpha, \beta\}$. Here is the reward function:

	α	β
1A	0	100
1B	0	200
2A	0	50
2B	0	400
3A	0	0
3B	0	600
T	0	0

As for the probability transition function, $P(T|T) = 1$, and $P(T|\beta) = 1$. Transition probabilities of non-terminal states for the α action are:

	2A	2B	3A	3B	T
1A	.4	.6			
1B	.6	.4			
2A			.4	.6	
2B			.6	.4	
3A					1
3B					1

Perform policy iteration on this Markov model, showing the state values, action values, and recommended policy after each iteration.

Initialize the policy $\pi(s)$ to β for each state s.
Use a discount factor of 1 (in other words, no discount).
In the policy iteration step, if $Q(s, a)$ is the same for both actions, default the action to β.
Do as many iterations necessary for the policy to converge.

Solution:

Iteration 1

	1	2	3
A	100	50	0
B	200	400	600

	1,α	1,β	2,α	2,β	3,α	3,β
A	260	100	360	50	0	0
B	190	200	240	400	0	600

	1	2	3
A	α	α	β
B	β	β	β

Iteration 2

	1	2	3
A	384	360	0
B	200	400	600

	1, α	1, β	2, α	2, β	3, α	3, β
A	384	100	360	50	0	0
B	376	200	240	400	0	600

	1	2	3
A	α	α	β
B	α	β	β

Iteration 3

	1	2	3
A	384	360	0
B	376	400	600

	1, α	1, β	2, α	2, β	3, α	3, β
A	384	100	360	50	0	0
B	376	200	240	400	0	600

	1	2	3
A	α	α	β
B	α	β	β

Policy converges.

Problem 8.5

The informational entropy of a probability distribution is defined as the - expected-value $\log(p(X))$. (Use a base 2 log)

a) Calculate the entropy of a fair coin toss (one that is heads half the time and tails the other half.

b) Calculate the entropy of a biased coin that is heads 75% of the time.

c) Explain in your own words what you think entropy measures and explain why the answer to (1) is greater than the answer to (2).

Solution:

a) $H(.5, .5) = -.5 \log .5 - .5 \log .5 = 1$ or equivalently $H(.5, .5) = \log |X| = \log 2 = 1$

b) $H(.75, .25) = -.75 \log .75 - .25 \log .25 = 0.8113$

c) Entropy measures the amount of information in a random variable or probability distribution. We have more information about the biased coin since we know that it is more likely that it lands on heads vs. tails. Thus the amount of information we learn by observing the biased coin is less than the amount of information we learn by observing the unbiased coin.

Problem 8.6

Prove that in policy iteration, the sequence of utilities $U^{\pi k}$ is non-decreasing. Namely, show that $U^{\pi k+1}(s) \geq U^{\pi k}(s)$ for all s.

Solution:

On the kth iteration of policy evaluation, for each state s,

$$\pi_{k+1}(s) = \arg\max_a \sum_{s'} (R(s, a, s') + \gamma U^{\pi_k}(s'))T(s, a, s')$$

For every state s:

$$U^{\pi_k}(s) \leq \sum_{s'} (R(s, \pi_{k+1}(s), s') + \gamma U^{\pi_k}(s'))T(s, \pi_{k+1}(s), s')$$

Thus, as we iterate over the states, the value we calculate for each state will be non-decreasing.

So,

$$\sum_{s'} (R(s, \pi_{k+1}(s), s') + \gamma U^{\pi_k}(s'))T(s, \pi_{k+1}(s), s') \leq U^{\pi_{k+1}}(s)$$

Thereby $U^{\pi_{k+1}}(s) \geq U^{\pi_k}(s)$ for all s, which is what we wished to show.