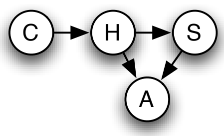
Project 3.2 Solutions

CS 188, Fall 2006



Question 1 (1)

(2) $P(A|C = c, S = s) = \alpha \sum_h P(A, h, c, s) = \alpha \sum_h P(c)P(h|c)P(s|h)P(A|h, s)$;

$$\alpha = \frac{1}{\sum_a \sum_h P(c)P(h|c)P(s|h)P(a|h, s)}$$

A	$\frac{1}{\alpha}P(A|c, s)$	$P(A|c, s)$
a	$0.3 \cdot 0.6 \cdot 0.9 \cdot 0.01 + 0.3 \cdot 0.1 \cdot 0.7 \cdot 0.5 + 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.2$	0.08
$\neg a$	$0.3 \cdot 0.6 \cdot 0.9 \cdot 0.99 + 0.3 \cdot 0.1 \cdot 0.7 \cdot 0.5 + 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.8$	0.92

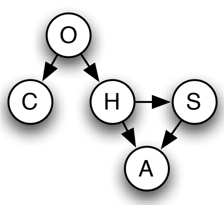
(3) $P(A|c, s) \approx \frac{1}{5}$ by counting

(4)

Sample	Weight $= P(c) \cdot P(s|h)$
$c, f, s, \neg a$	$0.3 \cdot 0.9$
$c, f, s, \neg a$	$0.3 \cdot 0.9$
$c, f, s, \neg a$	$0.3 \cdot 0.9$
c, p, s, a	$0.3 \cdot 0.3$
$c, v, s, \neg a$	$0.3 \cdot 0.7$

, so $P(A|c, s) \approx \frac{0.09}{1.11} = 0.08$

(5) The better estimate given by likelihood weighting indicates that these samples were probably generated by likelihood weighting. One could compute the likelihood of the samples under each scheme, but we did not expect this.

(6) 

, where O indicates whether the cat has an *owner*.

Question 2 a. A, C, E, F, G

b. A, C, D, E, F

c. A,B,E

d. A,D

Question 3

a. Under G, $P(x|y) = \alpha \sum_z P(x,y,z) = \alpha \sum_z P(x)P(y|x)P(z|y) = \alpha P(x)P(y|x) \sum_z P(z|y)$. Since $P(Z|y)$ is a distirbution over Z, $\sum_z P(z|y) = 1$. So, $P(x|y) = \alpha P(x)P(y|x)$.

Under G', $P(x|y) = \alpha P(x,y) = \alpha P(x)P(y|x)$, which is the same as under G.

In both calculations, α is the normalizing factor $\frac{1}{\sum_x P(x)P(y|x)}$.

b.
$$P(q_1,...,q_k|e_1,...,e_m) = \frac{\sum_{h_1} ... \sum_{h_p} \ P(q_1,...,q_k,e_1,...,e_m,h_1,...,h_p)}{\sum_{h_1} ... \sum_{h_p} \sum_{q_1} ... \sum_{q_k} \ P(q_1,...,q_k,e_1,...,e_m,h_1,...,h_p)}$$

c. Let $\pi(n)$ denote the parents of node n in the Bayes net. Then, we can rewrite

$$P(q_1,...,q_k,e_1,...,e_m,h_1,...,h_p) = \prod_{i=1}^{k} P(q_i|\pi(q_i)) \prod_{j=1}^{m} P(e_j|\pi(e_i)) \prod_{l=1}^{k} P(h_l|\pi(h_l))$$

which just follows from the semantics of Bayes nets. Notice that since h_1 is a leaf, none of the factors above contains h_1 except $P(h_1|\pi(h_1))$. So, we know that

$$\begin{aligned}
P(q_1,...,q_k|e_1,...,e_m) &= \alpha \sum_{h_1} ... \sum_{h_p} \ P(q_1,...,q_k,e_1,...,e_m,h_1,...,h_p) \\
&= \alpha \sum_{h_1} ... \sum_{h_p} \prod_{i=1}^{k} P(q_i|\pi(q_i)) \prod_{j=1}^{m} P(e_j|\pi(e_i)) \prod_{l=1}^{k} P(h_l|\pi(h_l)) \\
&= \alpha \sum_{h_2} ... \sum_{h_p} \prod_{i=1}^{k} P(q_i|\pi(q_i)) \prod_{j=1}^{m} P(e_j|\pi(e_i)) \prod_{l=2}^{k} P(h_l|\pi(h_l)) \sum_{h_1} P(h_1|\pi(h_1)) \\
&= \alpha \sum_{h_1} ... \sum_{h_p} \prod_{i=1}^{k} P(q_i|\pi(q_i)) \prod_{j=1}^{m} P(e_j|\pi(e_i)) \prod_{l=2}^{k} P(h_l|\pi(h_l))
\end{aligned}$$

Where the last step follows from the fact that $\sum_{h_1} P(h_1|\pi(h_1)) = 1$, much like in part (1). Note that the first line is just a restatement of part (2), where α is the normalizing factor (denominator).

d. If we successively remove hidden leaf nodes from the graph (which must be dangling), we will leave $P(q_1,...,q_k|e_1,...,e_m)$ unchanged by part (3). Let d be a dangling node. Then,n will eventually be pruned because all its descendants are also dangling. Before n is pruned, there will always be at least one dangling leaf because the graph is acyclic.