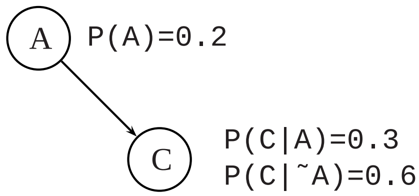
10-701/15-781 Machine Learning, Fall 2003

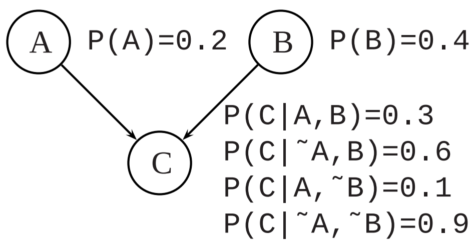
Homework 5 Solution

If you have questions, please contact Jiayong Zhang <zhangjy@cs.cmu.edu>.

1. (**40pts**, Evaluation) For (a)-(e), compute the following probabilities from the given Bayes nets. These examples have been designed so that none of the calculations should take you longer than a few minutes. If you find youself doing dozens of calculations on a question sit back and look for shortcuts.

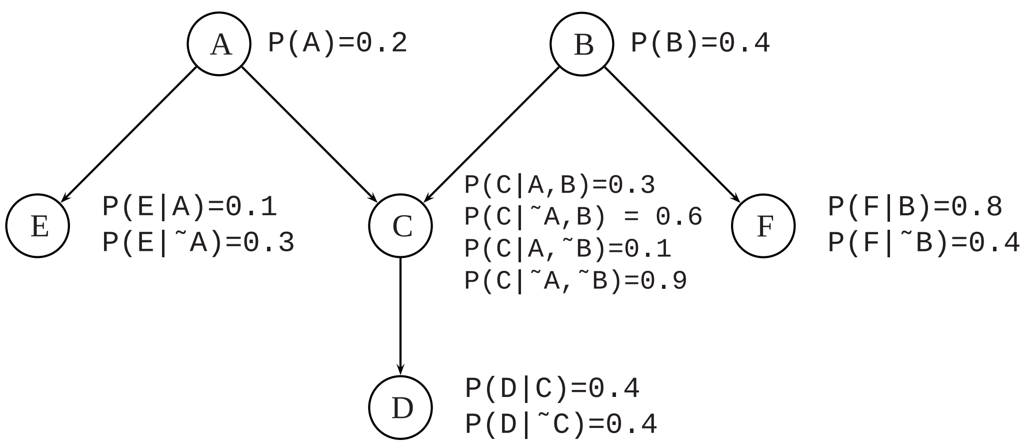


(a) $P(A|C) = \dfrac{P(A,C)}{P(A,C) + P(\sim A,C)} = \dfrac{1}{9}$

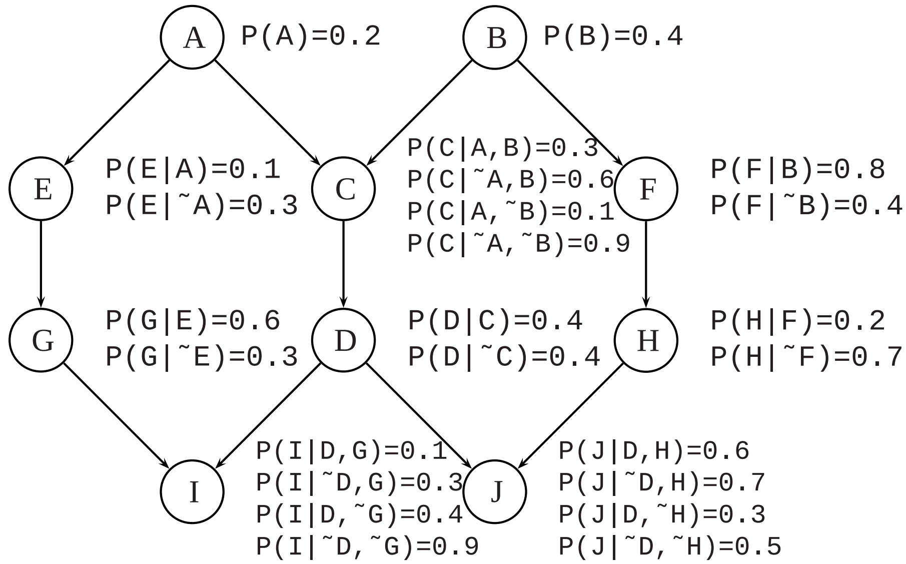


(b) $P(\sim A|B) = P(\sim A) = 0.8$

(c) $P(\sim A|B,\sim C) = \dfrac{P(\sim A,B,\sim C)}{P(\sim A,B,\sim C) + P(A,B,\sim C)} = \dfrac{16}{23} = 0.696$

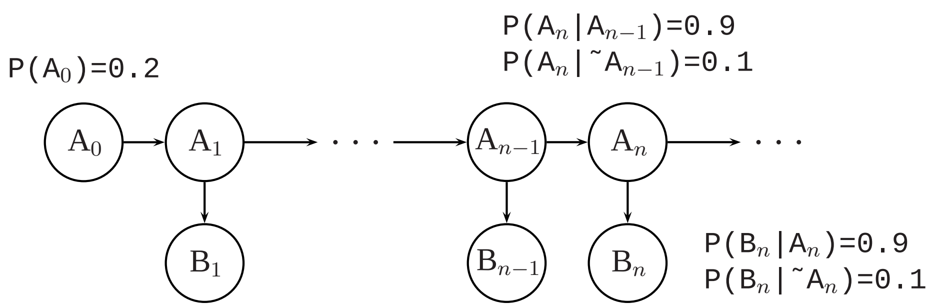


(d) $P(E|D) = P(E) = P(E,A) + P(E,\sim A) = 0.26$ (Link CD can be removed.)



(e) $P(\tilde{}G|\tilde{}J) = P(\tilde{}G) = P(\tilde{}G, E) + P(\tilde{}G, \tilde{}E) = 0.622$

(Remove link CD. Then G and J are d-separated.)



Let $q_n = P(A_n|B_1, B_2, \ldots, B_n)$.

(f) Compute q_n in terms of q_{n-1}.

$$\begin{aligned} P(A_n|B_{1:n-1}) &= P(A_n|A_{n-1})P(A_{n-1}|B_{1:n-1}) + P(A_n|\tilde{}A_{n-1})P(\tilde{}A_{n-1}|B_{1:n-1}) \\ &= 0.9q_{n-1} + 0.1(1 - q_{n-1}) = 0.8q_{n-1} + 0.1 \end{aligned}$$

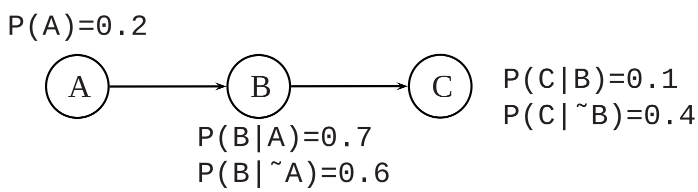
$$P(\tilde{}A_n|B_{1:n-1}) = 1 - P(A_n|B_{1:n-1}) = 0.9 - 0.8q_{n-1}$$

$$\begin{aligned} q_n = P(A_n|B_{1:n-1}, B_n) &= \frac{P(A_n, B_n|B_{1:n-1})}{P(A_n, B_n|B_{1:n-1}) + P(\tilde{}A_n, B_n|B_{1:n-1})} \\ &= \frac{P(B_n|A_n, B_{1:n-1})P(A_n|B_{1:n-1})}{P(B_n|A_n, B_{1:n-1})P(A_n|B_{1:n-1}) + P(B_n|\tilde{}A_n, B_{1:n-1})P(\tilde{}A_n|B_{1:n-1})} \\ &= \frac{0.9(0.8q_{n-1} + 0.1)}{0.9(0.8q_{n-1} + 0.1) + 0.1(0.9 - 0.8q_{n-1})} = \frac{0.72q_{n-1} + 0.09}{0.64q_{n-1} + 0.18} \end{aligned}$$

(g) What is $\lim_{n\to\infty} q_n$?

$q_n = q_{n-1} \Longrightarrow 0.64q^2 - 0.54q - 0.09 = 0 \Longrightarrow q \approx 0.9863.$

2. (**30 pts**, Likelihood Weighting) For this part, you will be writing a program to compute the approximate probabilities of events given a particular Bayes net. You are asked to use the approach of likelihood weighting described in your class notes. As an example, we can estimate $P(A|C)$ in the following Bayes net.



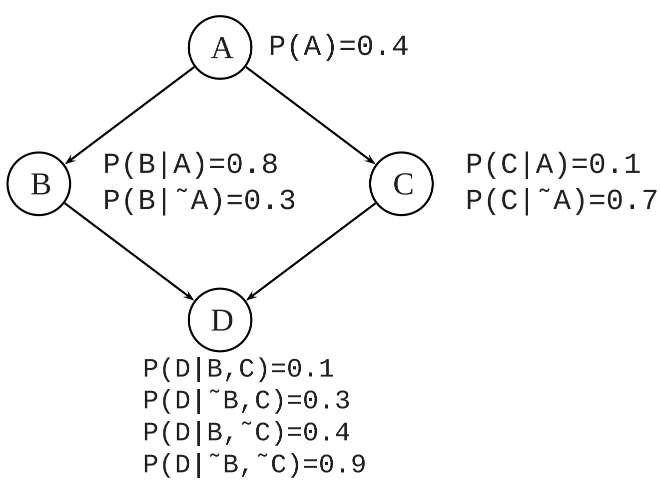
A C-code implentation `example.c` is provided. The program can be compiled from `gcc` on UNIX/LINUX machines with

```
gcc -o example example.c -lm
```

and then runs as

```
./example <num_samples>
```

Now consider the following Bayes net.



(a) Hand in source code similar to the example that computes $P(B|C)$ using likelihood weighting. You may use `example.c` and modify for your own answer if you wish. Or you can use any other language you prefer.

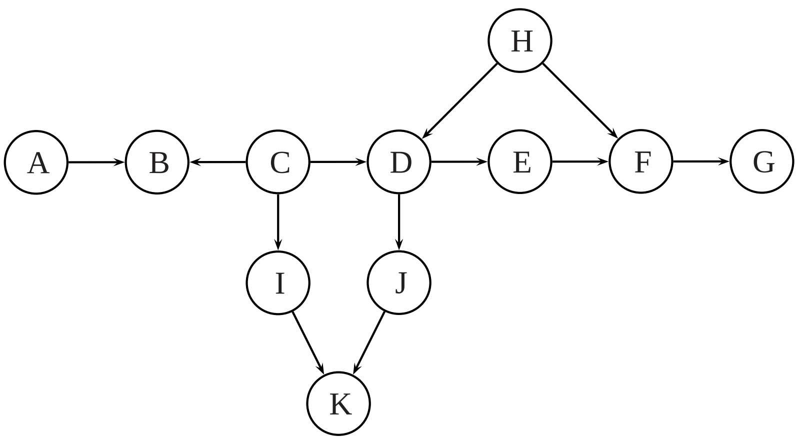
Critical modification:

```
a = coin_toss(0.4);
b = (a==0)?coin_toss(0.3):coin_toss(0.8);

if (a==0) weight *= 0.7;
else weight *= 0.1;

if (b==1)  sum_w1 += weight;
sum_w2 += weight;
```

(b) Plot the average absolute estimation error versus the number of samples.

Note that the question asks for *average* error, so multiple runs for each number of samples are required.

3. (**30 pts**, D-separation) Using the given Bayes network, for each of the following statements indicate whether it is true or false.



(a) I<A,B,C>

FALSE. The only path ABC is not blocked because the arcs are head-to-head and B is in the evidence set.

(b) I<D,E,F>

FALSE. The path DHF is not blocked because the arcs are tail-to-tail and H is not in the evidence set.

(c) I<A,{},D>

TRUE. Any path from A to D is blocked by B which has head-to-head arcs and is not in the evidence set.

(d) I<B,{},I>

FALSE. The path BCI is not blocked because C is not in the evidence set.

(e) `I<B,D,J>`

TRUE. The only paths are BCDJ and BCIKJ. The path BCDJ is blocked by D, because it has tail-to-head arcs and is in the evidence set. The path BCIKJ is blocked by K, because it has head-to-head arcs and is not in the evidence set. It has not descendents, so its descendents are not in the evidence set either.

(f) `I<C,{G},H>`

FALSE. The path CDH is unblocked because D has head-to-head arcs and its descendent G is in the evidence set.

(g) `I<I,J,H>`

FALSE. ICDH is not blocked because D has head-to-head arcs and its descendent J is in the evidence set.

(h) `I<A,{B,E},G>`

FALSE. ABCDHFG is unblocked because 1) B has head-to-head arcs and is in the evidence set, and 2) D has head-to-head arcs but its descendent E is in the evidence set.

(i) `I<K,{},G>`

FALSE. KJDEFG is unblocked because it only contains tail-to-tail and head-to-tail arcs and no variable in the path is in the evidence set.

(j) `I<C,{},H>`

TRUE. The path CDH is blocked by D, the path CDEFH is blocked by F, and the path CIKJDH is blocked by K.