

Names: Will Chen, Noel Declaro, Afsana Rahman

Overview: This document contains answers to Problems 1-4 in Assignment 2.

Question 1 (10 points): Consider the following Bayesian network, where variables A through E are all boolean valued:

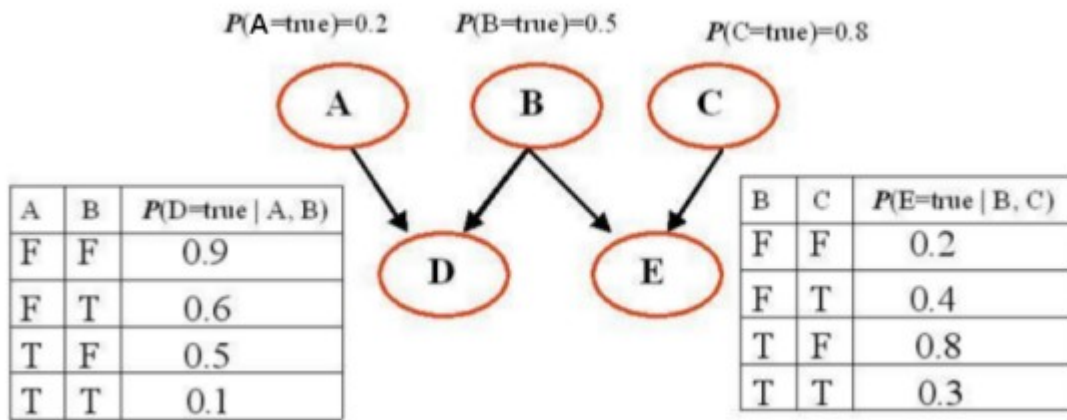


Figure 1

For the sake of shorthand, assume for any random variable A, B, C, D, E in this network, its corresponding lowercase letter refers to its true value being *true* and its negation being *false* (e.g., $a \iff A = \text{true}$, $\neg a \iff A = \text{false}$).

(a) **Probability of all true.** We want to find $P(a, b, c, d, e)$.

- $P(a) = 0.2$
- $P(b) = 0.5$
- $P(c) = 0.8$
- $P(d|a, b) = 0.1$
- $P(e|b, c) = 0.3$

$$\begin{aligned}
 P(a, b, c, d, e) &= P(a) \times P(b) \times P(c) \times P(d|a, b) \times P(e|b, c) \\
 &= 0.2 \times 0.5 \times 0.8 \times 0.1 \times 0.3 \\
 &= 0.0024
 \end{aligned}$$

(b) **Probability of all false.** We want to find $P(\neg a, \neg b, \neg c, \neg d, \neg e)$.

- $P(\neg a) = 0.8$

- $P(\neg b) = 0.5$
- $P(\neg c) = 0.2$
- $P(\neg d|\neg a, \neg b) = 0.1$
- $P(\neg e|\neg b, \neg c) = 0.3$

$$\begin{aligned}
 P(\neg a, \neg b, \neg c, \neg d, \neg e) &= P(\neg a) \times P(\neg b) \times P(\neg c) \times P(\neg d|\neg a, \neg b) \times P(\neg e|\neg b, \neg c) \\
 &= 0.8 \times 0.5 \times 0.2 \times 0.9 \times 0.2 \\
 &= 0.0064
 \end{aligned}$$

(c) **Probability A is false given rest are true.**

- $P(\neg a) = 0.8$
- $P(b) = 0.5$
- $P(c) = 0.8$
- $P(d|\neg a, b) = 0.6$
- $P(e|b, c) = 0.3$

$$\begin{aligned}
 P(\neg a, b, c, d, e) &= P(\neg a) \times P(b) \times P(c) \times P(d|\neg a, b) \times P(e|b, c) \\
 &= 0.8 \times 0.5 \times 0.8 \times 0.6 \times 0.3 \\
 &= 0.0576
 \end{aligned}$$

Question 2 (10 points):

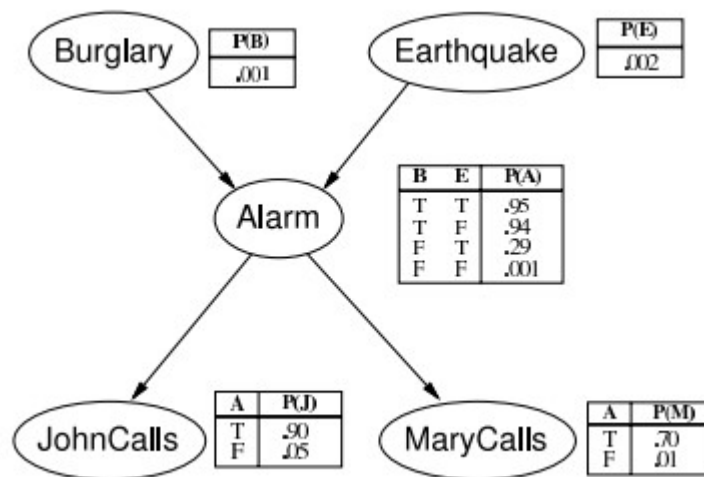


Figure 2

- (a) **Variable elimination.** Given the query $P(\text{Burglary} | \text{JohnCalls} = \text{true}, \text{MaryCalls} = \text{true})$, we can rewrite as:

$$P(B|j, m) = \alpha P(B) \sum_e P(e) \sum_a P(a|B, e) P(j|a) P(m|a) \quad (1)$$

We can then begin to eliminate variables, starting with the variables including a :

$$P(B|j, m) = \alpha P(B) \sum_e P(e) \begin{pmatrix} 0.598582 & 0.18655 \\ 0.59223 & 0.0011295 \end{pmatrix}$$

And then "eliminate" $\sum_e P(e)$:

$$P(B|j, m) = \alpha P(B) \begin{pmatrix} 0.0002 * 0.598525 + 0.998 * 0.59223 \\ 0.0002 * 0.183055 + 0.998 * 0.0011295 \end{pmatrix}$$

And then "eliminate" $P(B)$:

$$P(B|j, m) = \alpha \begin{pmatrix} 0.001 * 0.59224259 \\ 0.999 * 0.001491858 \end{pmatrix}$$

And then normalize. So we have $P(B|j, m) = \langle 0.284, 0.716 \rangle$.

- (b) **Number of operations.**

- **Variable elimination.** In this problem, computing takes two sets of 3 operations. Since there are 11 total operations in burglary, we get $2 * 11 + 3 = 25$ operations.
- **Enumeration.** (As taken from the lecture slides) The expression is computed by adding four terms, each computed by multiplying five numbers. So there is a total of 40 operations (since time is doubled due to computation of α).

- (c) **Chain Bayesian network.**

- **Variable elimination.** There are $n - 2$ variables to sum over with $n - 1$ multiplications each.

$$P(X_1 | X_n = \text{true}) = \alpha \sum_{X_2 \dots X_{n-1}} P(X_n = \text{true} | X_{n-1}) P(X_{n-1} | X_{n-2}) \dots P(X_2 | X_1) P(X_1)$$

$$P(X_1 | X_n = \text{true}) = \alpha \sum_{X_{n-1}} P(X_n = \text{true} | X_{n-1}) \sum_{X_{n-2}} P(X_{n-1} | X_{n-2}) \dots \sum_{X_2} P(X_3 | X_2)$$

There are $n - 2$ summations on each; it has running time $O(n)$.

- **Enumeration.** Two binary search trees can be used for each value of x_1 with depths $n - 2$, giving total work done to be $O(2n)$.

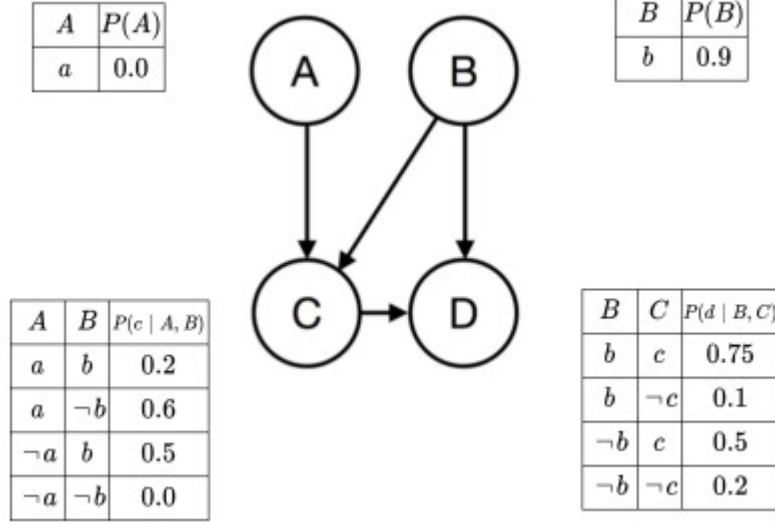


Figure 3

Question 3 (15 points): Please note the code used in this question was included in the zip file of our submission.

(a) **Enumeration.**

- $P(d|c)$. To find this probability, we'll introduce α and sum over a and b .

$$\begin{aligned}
 P(d|c) &= \alpha P(d, c) = \alpha P(a, b, c, d) \\
 &= \alpha \sum_a \sum_b P(a)P(b)P(c|a, b)P(d|b, c) \\
 &= \alpha \sum_a P(a) \sum_b P(b)P(c|a, b)P(d|b, c) \\
 &= \alpha < 0.3375, 0.1125 > = < 0.75, 0.25 > \\
 \text{This means } P(d|c) &= 0.75.
 \end{aligned}$$

- $P(b|c)$. To find this probability, we'll introduce α and sum over a and d .

$$\begin{aligned}
 P(b|c) &= \alpha P(b, c) = \alpha P(a, b, c, d) \\
 &= \alpha \sum_a \sum_d P(a)P(b)P(c|a, b)P(d|b, c) \\
 &= \alpha P(b) \sum_a P(a)P(c|a, b) \sum_d P(d|b, c) \\
 &= \alpha < 0.45, 0 > = < 1, 0 > \\
 \text{This means } P(b|c) &= 1.
 \end{aligned}$$

- $P(d|\neg a, b)$. To find this probability, we'll introduce α and sum over c .

$$\begin{aligned}
 P(d|\neg a, b) &= \alpha P(d, a, b) = \alpha P(a, b, c, d) \\
 &= \alpha \sum_c P(a)P(b)P(c|a, b)P(d|b, c) \\
 &= \alpha P(a)P(b) \sum_a P(c|a, b) \sum_d P(d|b, c) \\
 &= \alpha < 0.3825, 0.5157 > = < 0.425, 0.575 > \\
 \text{This means } P(d|\neg a, b) &= 0.425.
 \end{aligned}$$

- (b) **Likelihood and rejection sampling.** The results of 1000 samples are shown in Figure 4. The approximation using rejection sampling was fairly accurate to the true probabilities. The approximation using likelihood was similar.

```

rejsampout.txt
Samples passed for P(d|c): 444, samples true: 332
NORMALIZATION = 0.747747747748

Samples passed for P(b|c): 444, samples true: 444
NORMALIZATION = 1.0

Samples passed for P(d|nota, b): 904, samples true: 384
NORMALIZATION = 0.424778761062

```

(a)

```

likelihoodout.txt
Total weight for P(d|c): 447.5, sample true: 339.0
NORMALIZATION = 0.757541899441

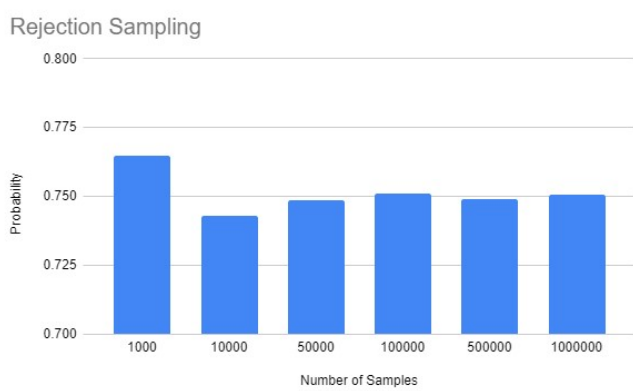
Total weight for P(b|c): 447.5, sample true: 447.5
NORMALIZATION = 1.0

Total weight for P(d|nota, b): 900.0, sample true: 397.8
NORMALIZATION = 0.442

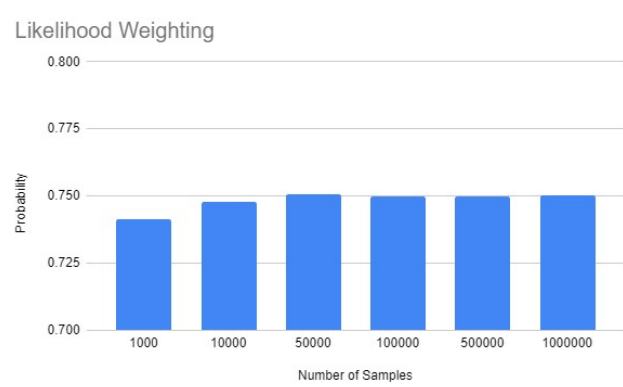
```

(b)

Figure 4: Results of rejection sampling and likelihood weighting, respectively, for 1000 samples.



(a)



(b)

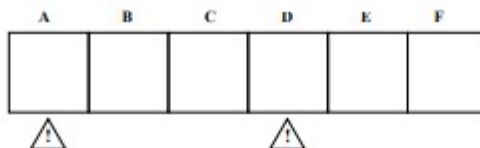
Figure 5: Plots of rejection sampling and likelihood weighting, respectively, for increasing samples.

- (c) **Number of samples.** The plot of probability versus number of samples is shown in Figure 5. The number of samples was changed manually in the code from the previous part. The graphs were also created manually. The likelihood weighting was shown to converge much faster than the rejection sampling.

Question 4 (15 points):

- (a) **Filtering.** We want to find the probability distribution over the rover's position on day 3, given the observations ($E_1 = hot, E_2 = cold, E_3 = cold$). It is important to note that the rover does not misread temperature, and only moves 1 mile east per day or stays in place for the day.

This means that the minimum distance the rover could move is 0 miles, meaning it stays in position A, and the maximum distance is 2 miles to position C (since there is no movement on day 1). This means that it is automatically impossible for the rover to be at positions D, E, or F, or in other words, the probability of the rover being positioned anywhere here on day 3 is 0.



The rover was heavily damaged in the fall, and as a result, most of its sensors are broken. The only ones still functioning are its thermometers, which register only two levels: *hot* and *cold*. The rover sends back evidence $E = \text{hot}$ when it is at a volcanic vent (A and D), and $E = \text{cold}$ otherwise. There is no chance of a mistaken reading. The rover fell into the gorge at position A on day 1, so $X_1 = A$. Let the rover's position on day t be X_t , where $X_t \in \{A, B, C, D, E, F\}$ on any day. The rover is still executing its original programming, trying to move 1 mile east (i.e., right, towards F) every day. Because of the damage, however, it only moves east with probability 0.80 on any one day, and it stays in place with probability 0.20 on any one day. Your job is to figure out where the rover is, so that you can dispatch your rescue-bot.

Figure 6

Now, since $E_2 = \text{cold}$, and we know the rover does not make false temperature reports, and the rover only moves east, the rover must be at position B on day 2 (making the probability of the rover being at position A on day 3 therefore 0). Since $E_3 = \text{cold}$ as well, there is a chance the rover could have either stayed in place, or moved to position C, meaning the probability distribution on day 3 for the rover would be:

$$\text{DAY 3: } A = 0, B = 0.2, C = 0.8, D = 0, E = 0, F = 0$$

- (b) **Smoothing.** Like above, we want to find the probability distribution of the rover's position, except now on day 2. As explained above, since the rover does not make false temperature reports, and we have recorded $E_2 = \text{cold}$, then we know the rover position cannot be A or D. We also know that since the rover only moves east, 1 mile per day maximum, it cannot have gone past position C (non-inclusive) before day 3. Therefore, the rover must be at position B on day 2 given the evidence, i.e.,

$$\text{DAY 2: } A = 0, B = 1, C = 0, D = 0, E = 0, F = 0$$

- (c) **Most Likely Explanation.** The possible sequences that may have occurred on the first three days are either:

- The rover starts at A, moves to B on day 2, and stay on B for day 3. OR
- The rover starts at A, moves to B on day 2, and moves to C on day 3.

The only difference between the two sequences is that the rover either moves from or stays at position B on day 3. Since the former is more likely (0.8 versus 0.2 chance), the more likely *overall* sequence is the one where the rover keeps moving east, i.e., position sequence A, B, C for days 1,2,3, respectively.

- (d) **Prediction.** Given the original sequence of temperature readings, $E_1 = \text{hot}, E_2 = \text{cold}, E_3 = \text{cold}$, we want to find the probability of the next sequence $E_4 = \text{hot}, E_5 =$

hot, $E_6 = \text{cold}$. The only possible way for the readings to be accurately $E_3 = \text{cold}$, $E_4 = \text{hot}$ is for the rover to be at position C on day 3, and then move to position D on day 4. Similarly, the *only* possibly way for the rover to read "hot" two days in a row is if it stays on D for two days. *Similarly*, the *only* way for the rover to move from hot to cold is for the rover to be at position D and then move to E. In summary, the only possible location sequence for all the evidence to be true is (for days 1-6 in chronological order) A, B, C, D, D, E.

As determined in part (a), the probability of the rover being at position C on day 3 is 0.8; the chance of moving east on day 4 is also 0.8; the chance of it staying at its new location (D) on day 5 is 0.2; and the chance of it moving east on day 6 is 0.8. Therefore, the probability of the new sequence occurring given the previous one is $0.8 * 0.8 * 0.2 * 0.8 = 0.1024$.

(e) **Prediction.** Basing off part (a)'s probability distribution:

$$\text{DAY 3: } A = 0, B = 0.2, C = 0.8, D = 0, E = 0, F = 0$$

we can build predictions for the next day based the on the information we are given. Since temperature is not taken into account, we only need to regard the probabilities of the rover's movement.

Regarding day 4, if we assume that the rover is at position B, there is a 0.2 chance it stays there, and 0.8 percent chance it moves east to position C. Since the probability of being at position B was already 0.2 on day 3, the chance of staying there would be $0.2 * 0.2 = 0.04$, and the probability of moving from there would be $0.2 * 0.8 = 0.16$. The same logic applies to position C: the chance of being at position C on day 3 is 0.8, there is 0.2 chance of staying there, and 0.8 chance of moving east to position D. Therefore, if assuming we start at position C, the probability of staying there is $0.8 * 0.2 = 0.16$ and the probability of moving to position D is $0.8 * 0.8 = 0.64$. In summary:

$$\text{DAY 4: } A = 0, B = 0.04, C = 0.16 + 0.16 = 0.32, D = 0.64, E = 0, F = 0$$

which is normalized. We repeat this process for day 5:

- **Position B.** There is a 0.04 chance of the rover being here on day 4, and 0.2 chance of it staying for day 5, so the rover has a $0.04 * 0.2 = 0.008$ chance overall of being there on day 5.
- **Position C.** There is a 0.04 chance of the rover being at position B on day 4 with 0.8 chance of moving to position C, meaning this course of action has $0.04 * 0.8 = 0.032$ chance of happening. There is also a 0.32 chance the rover is already there on day 4, and 0.2 chance it may stay there for day 5, meaning there is $0.32 * 0.2 = 0.064$ chance of this course of actions occurring. This means the overall chance of the rover being here on day 5 is $0.032 + 0.064 = 0.096$.
- **Position D.** There is a 0.32 chance of the rover being at position C on day 4 with 0.8 chance of moving to position D, meaning this course of action has $0.32 * 0.8 = 0.256$

chance of happening. There is also a 0.64 chance the rover is already there on day 4, and 0.2 chance it may stay there for day 5, meaning there is $0.64 * 0.2 = 0.128$ chance of this course of actions occurring. This means the overall chance of the rover being here on day 5 is $0.256 + 0.128 = 0.384$.

- **Position E.** This is the furthest possible position the rover can go, as it has not been explored yet. There is a 0.64 chance the rover is at position D at day 4, and 0.8 chance the rover will move here on day 5, so the rover has $0.64 * 0.8 = 0.512$ chance of being here on day 5.

With normalization of the values determined above:

$$\text{DAY 5: } A = 0, B = 0.008, C = 0.096, D = 0.384, E = 0.512, F = 0$$