

Probabilistic Reasoning  
CS:440 Assignment 2

Christopher Naporlee - cmn134  
Michael Nelli - mnr73  
Timothy Walker - tpw32

April 2022

## Question 1

a)  $P(A, B, C, D, E)$

$$\begin{aligned} &= P(D|A, B)P(E|B, C)P(A)P(B)P(C) \\ &= (0.1)(0.3)(0.2)(0.5)(0.8) \\ &= 0.0024 \end{aligned}$$

b)  $P(\neg A, \neg B, \neg C, \neg D, \neg E)$

$$\begin{aligned} &= P(\neg D|\neg A, \neg B)P(\neg E|\neg B, \neg C)P(\neg A)P(\neg B)P(\neg C) \\ &= (1 - 0.9)(1 - 0.2)(1 - 0.2)(1 - 0.5)(1 - 0.8) \\ &= (0.1)(0.8)(0.8)(0.5)(0.2) \\ &= 0.0064 \end{aligned}$$

c)  $P(\neg A|B, C, D, E)$

$$\begin{aligned} &= P(\neg A, B, C, D, E) \\ &= P(\neg A)P(B)P(C) \sum_D P(D|A, B) \sum_E P(E|B, C) \end{aligned}$$

At this point we know that B and C are set to be true so we can easily derive the probability values for those. Also we can easily find E to be 0.3, because B, C are true. Finally, A is false and B is true thus D can only be 0.6.

$$= (0.8)(0.5)(0.8)(0.3)(0.6) = 0.0576$$

## Question 2

a)  $P(\text{Burglary} \mid \text{JohnCalls} = \text{True}, \text{MaryCalls} = \text{True})$

$$\begin{aligned}
 P(B|j, m) &= \alpha P(B, j, m) \\
 &= \alpha \sum_e \sum_a P(B, j, m, e, a) \\
 &= \alpha \sum_e \sum_a P(B)P(e)P(A|B, e)P(j|a)P(m|a) \\
 &= \alpha P(B) \sum_e P(e) \sum_a P(a|B, e)P(j|a)P(m|a) \\
 &= \alpha \underbrace{P(B)}_{f_1(B)} \underbrace{\sum_e P(e)}_{f_2(E)} \underbrace{\sum_a P(a|B, e)}_{f_3(A, B, E)} \underbrace{P(j|a)}_{f_4(A)} \underbrace{P(m|a)}_{f_5(A)} \\
 f_6(B, E) &= \sum_a f_3(A, B, E) \times f_4(A) \times f_5(A) \\
 &= (f_3(a, B, E) \times f_4(a) \times f_5(a)) + (f_3(\neg a, B, E) \times f_4(\neg a) \times f_5(\neg a)) \\
 &= \begin{pmatrix} 0.95 & 0.29 \\ 0.94 & 0.001 \end{pmatrix} \times (0.9) \times (0.7) + \begin{pmatrix} 0.05 & 0.71 \\ 0.06 & 0.999 \end{pmatrix} + (0.05) \times (0.01) \\
 &= \begin{bmatrix} 0.598525 & 0.183055 \\ 0.592235 & 0.0011295 \end{bmatrix}
 \end{aligned}$$

Simplifying:

$$\begin{aligned}
 P(B|j, m) &= \alpha f_1(B) \times \sum_e f_2(E) \times f_6(B, E) \\
 f_7(B) &= \sum_e f_2(E) \times f_6(B, E) \\
 f_7(B) &= f_2(e) \times f_6(B, e) + f_2(\neg e) \times f_6(B, \neg e) \\
 &= ((.002) \times \begin{bmatrix} 0.598582 \\ 0.183055 \end{bmatrix}) + ((.998) \times \begin{bmatrix} 0.592235 \\ 0.0011295 \end{bmatrix}) = \begin{bmatrix} 0.59224 \\ 0.00149 \end{bmatrix} \\
 P(B|j, m) &= \alpha f_1(B) \times f_7(B) \\
 &= \alpha \begin{bmatrix} 0.001 \\ 0.999 \end{bmatrix} \times \begin{bmatrix} 0.59224 \\ 0.00149 \end{bmatrix} \\
 &\approx \begin{bmatrix} 0.284 \\ 0.716 \end{bmatrix}
 \end{aligned}$$

b) In our final equation we had  $P(B|j, m) = \alpha f_1(B) \times f_7(B)$ .

$$f_7(B) = f_2(e) \times f_6(B, e) + f_2(\neg e) \times f_6(B, \neg e)$$

$$f_6(B, E) = (f_3(a, B, E) \times f_4(a) \times f_5(a)) + (f_3(\neg a, B, E) \times f_4(\neg a) \times f_5(\neg a))$$

From  $f_6(B, E)$  we got:

$$f_7(B, E) = \left( \begin{bmatrix} 0.95 & 0.29 \\ 0.94 & 0.001 \end{bmatrix} \times (0.9) \times (0.7) \right) + \left( \begin{bmatrix} 0.05 & 0.71 \\ 0.06 & 0.999 \end{bmatrix} \times (0.05) \times (0.01) \right)$$

Which ends up resulting in 10 multiplications and 4 additions. This is because these operations are **pointwise**.

Next, from  $f_7(B)$  we got:

$$= ((.002) \times \begin{bmatrix} 0.598582 \\ 0.183055 \end{bmatrix}) + ((.998) \times \begin{bmatrix} 0.592235 \\ 0.0011295 \end{bmatrix})$$

which is another 4 multiplications and 2 additions. Finally our last step being:

$$= \alpha \begin{bmatrix} 0.001 \\ 0.999 \end{bmatrix} \times \begin{bmatrix} 0.59224 \\ 0.00149 \end{bmatrix}$$

To solve this part: 2 more pointwise multiplications. Followed by 1 addition and 1 division to solve  $\alpha$ . Finally doing 1 division to normalize. Bringing us to a grand total of 16 multiplications, 7 additions, and 2 divisions or 25 total operations.

As for enumeration we must:

$$P(b|j, m) = \frac{P(b, j, m)}{P(b, j, m) + P(\neg b, j, m)}$$

Where

$$P(b|j, m) = \alpha P(b) \sum_e \sum_a P(e) P(a|b, e) P(j|a) P(m|a)$$

In enumeration, we must do 3 multiplications over 2 values for each variable we are iterating over  $(e, a)$ . Thus,  $3 * 4 = 12$  multiplications with 3 additions as part of the summations. Finally, one final multiplication using  $P(b)$  to get the result. Leaving us with 13 multiplications and 3 additions.

However, we must calculate  $P(\neg b, j, m)$  in order to get the alpha. Thus: we get 16 operations + another 16 operations for 32 operations. Then another addition and division to apply our normalization.

Leaving us with a grand total of 34 operations to solve enumeration.

Comparing: variable elimination does  $(34 - 25) = 9$  less operations.

- c) First solving  $P(X_1|X_n = \text{true})$  using enumeration. We can represent the equation like:

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

In words, we must account for all the hidden variables in the bayesian network. In this problem, there are  $n - 1$  hidden variables ( $X_2 \dots X_{n-1}$ ). Therefore, we must take into account  $2^{n-1}$  entries. Thus, our total work to be done using enumeration will be  $O(2^{n-1})$  or more simply  $O(2^n)$ .

Next solving  $P(X_1|X_n = \text{true})$  using variable elimination. We have:

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

Variable elimination works by summing out the hidden variables that do not depend on the query:

$$\begin{aligned} P(X_1|X_n) &= P(X_1) \sum_{X_2} P(X_2|X_1) \sum_{X_3} P(X_3|X_2) \dots \sum_{X_{n-2}} P(X_{n-3}|X_{n-2}) f_1(X_{n-1}) \\ &= P(X_1) \sum_{X_2} P(X_2|X_1) \sum_{X_3} P(X_3|X_2) \dots \sum_{X_{n-3}} P(X_{n-4}|X_{n-3}) f_2(X_{n-2}) \\ &\quad \dots \end{aligned}$$

Therefore, using this method, we are able to not repeat work already done unlike variable enumeration. Further, since this network is a singly linked chain we can reference one of it's properties. "The time and space complexity of exact inference in poly-trees is linear in the size of the network" (Bayesian Networks and Exact Inference Notes, Pg. 5). Therefore, the amount of work done using this method will be  $O(n)$ .

### Question 3

- a) •  $P(d|c)$

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

From our coded sample tests of  $P(d|c)$  we got:

Rejection Sampling:  $< 0.74413, 0.25586 >$  with % Error  $\approx 0.80\%$

Weighted Sampling:  $< 0.74606, 0.25393 >$  with % Error  $\approx 0.53\%$

- $P(b|c)$

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

From our coded sample tests of  $P(b|c)$  we got:

Rejection Sampling:  $< 1, 0 >$  with % Error =  $0\%$

Weighted Sampling:  $< 1, 0 >$  with % Error =  $0\%$

- $P(d|\neg a, b)$

$$\begin{aligned}
 P(d|\neg a, b) &= \alpha P(d, \neg a, b) \\
 &= \alpha \sum_c P(a)P(b)P(c|a, b)P(d|b, c) \\
 &= \alpha P(a)P(b) \sum_c P(c|a, b)P(d|b, c) \\
 &= \alpha < 0.3825, 0.5175 > \\
 &= < 0.425, 0.575 >
 \end{aligned}$$

From our coded sample tests of  $P(d|\neg a, b)$  we got:

Rejection Sampling:  $< 0.42200, 0.57799 >$  with % Error  $\approx 0.71\%$

Weighted Sampling:  $< 0.43604, 0.55295 >$  with % Error  $\approx 2.50\%$

From our results we can see that our tests are very much similar to our expected probability.

b) Sampling Graphs on  $P(d|c)$

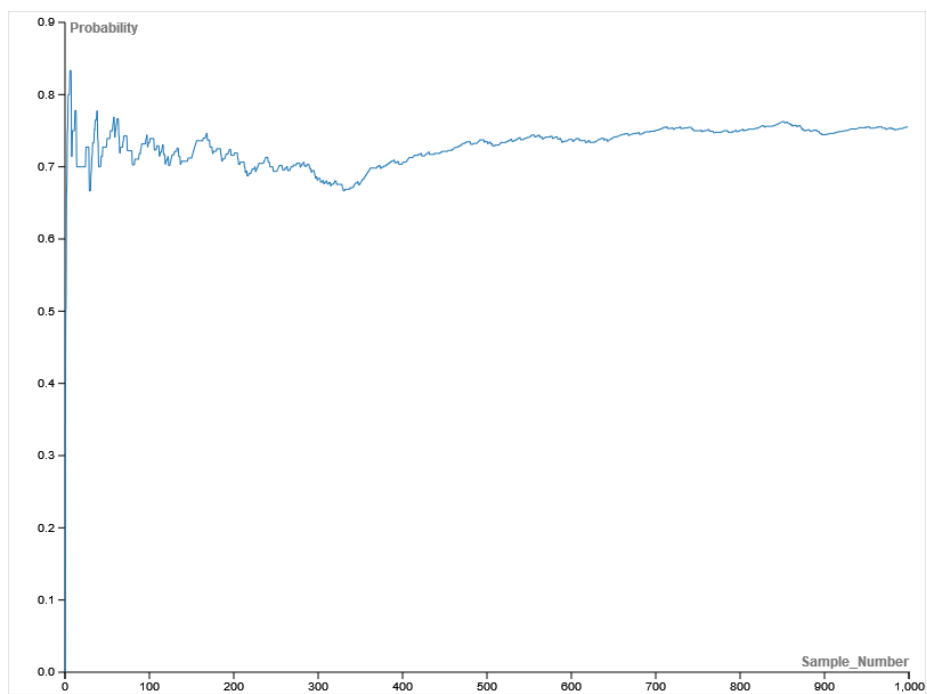


Figure 1: Rejection Sampling on  $P(d|c)$

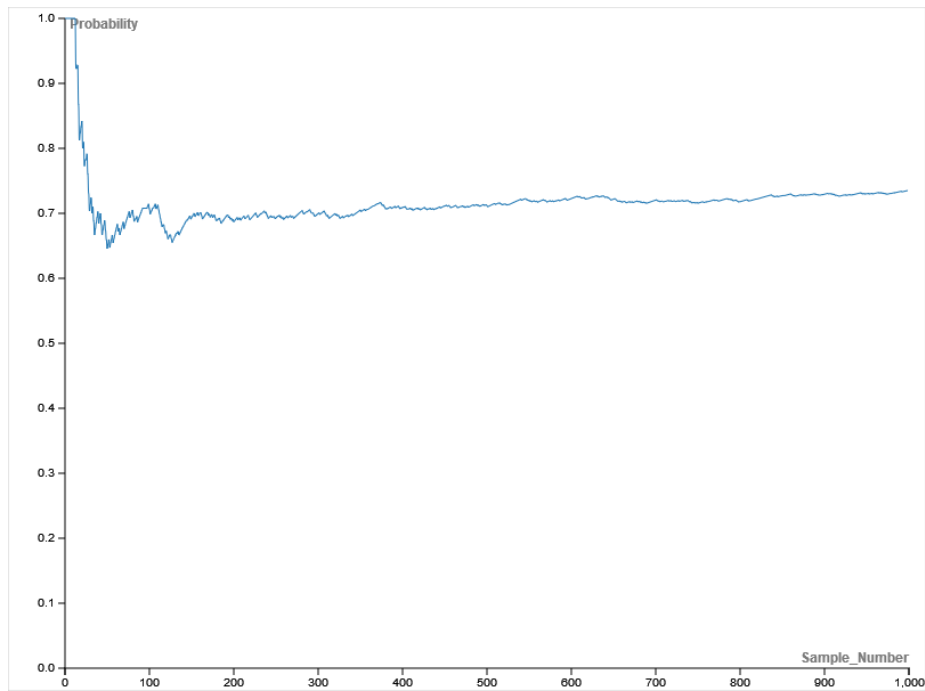


Figure 2: Weighted Sampling on  $P(d|c)$

As you can see, the weighted sampling graph starts out not quite close to the expected probability, but levels out much quicker. As opposed to the rejection graph which starts closer to the expected probability, but fluctuates a lot more. This is because (if you look close enough) there are samples being rejected which cause it to take much longer to stabilize at our expected probability.



c) Our formulated query we used was:  $P(d|-a, b, \neg c)$

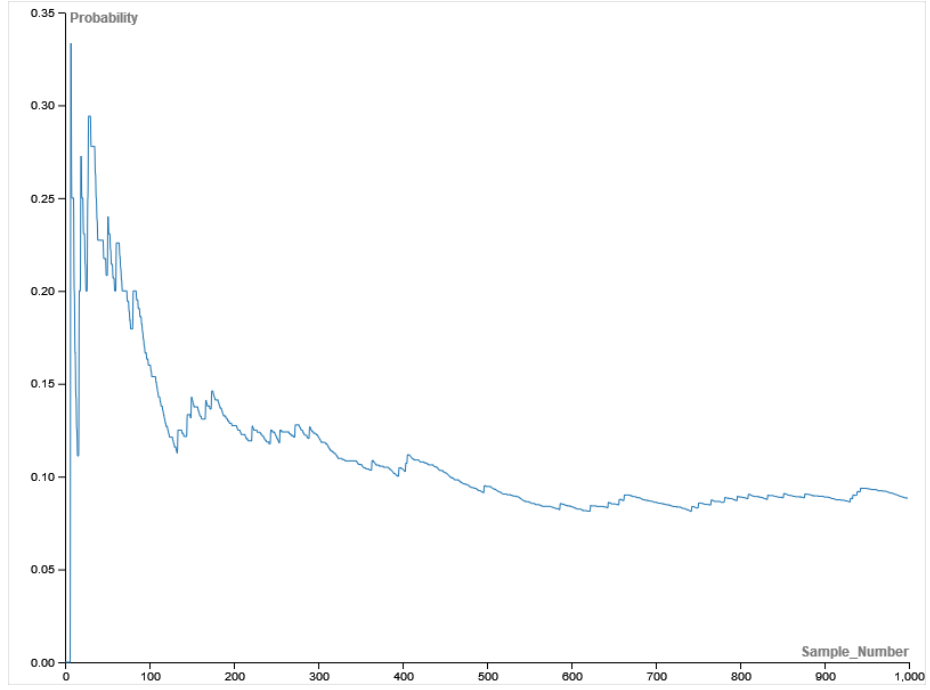


Figure 3: Rejection Sampling on  $P(d|-a, b, \neg c)$

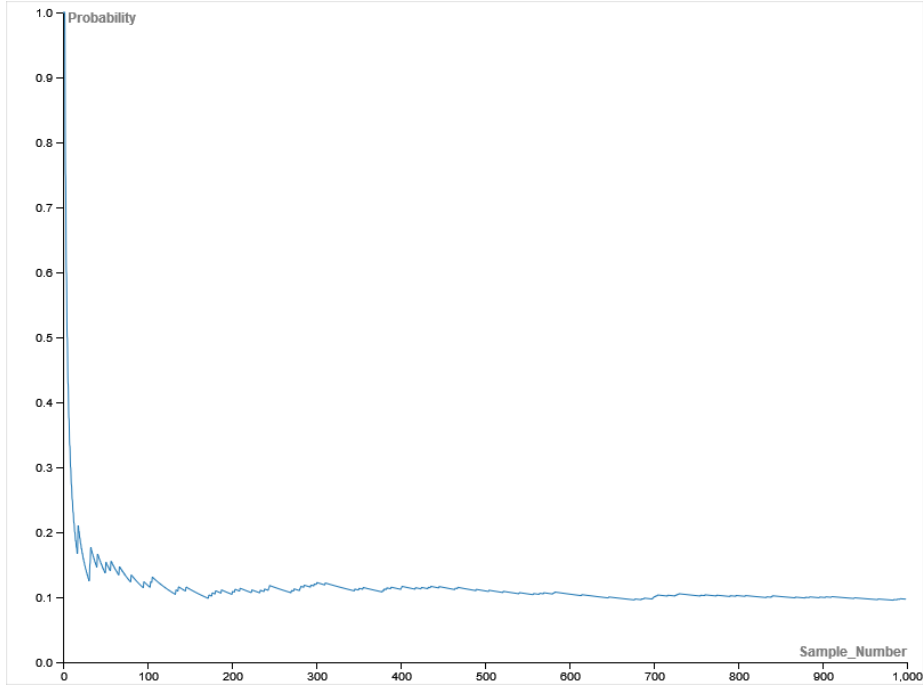


Figure 4: Weighted Sampling on  $P(d|-a, b, \neg c)$

From these graphs, it is apparent that rejection sampling was notably worse. This is due to rejection sampling's main pitfall - the fact that it will throw away too many samples. Given large amounts of evidence variables, as we have in our query of  $\neg a, b, \neg c$ , rejection sampling will throw out a large amount of samples as it struggles to find the correct atomic event to validate as a matching sample.

Knowing this, rejection sampling will need a much larger set of samples in order to reach the expected probability output. As opposed to weighting, that uses all the samples, it levels out much nicer and more quickly around 400 samples. For rejection sampling, it has still yet to level out even at 400 samples.

## Question 4

a) Given:  $X_1 = A$  Thus,  $P(X_1|hot_1) = 1$

$$\begin{aligned}
 & P(X_2|hot_1, cold_2) \\
 &= \alpha P(cold_2|X_2) \sum_{X_1} P(X_2|X_1)P(X_1|hot_1) \\
 &= \alpha < 0, 1, 1, 0, 1, 1 > \times < 0.2, 0.8, 0, 0, 0, 0 > \\
 &= \alpha < 0, 0.8, 0, 0, 0, 0 > \\
 &= < 0, 1, 0, 0, 0, 0 >
 \end{aligned}$$

This answer makes sense, as if we were to start on A on day 1 and day 2 read cold we must have moved from A to B (a cold tile). Continuing...

$$\begin{aligned}
 & P(X_3|hot_1, cold_2, cold_3) \\
 &= \alpha P(cold_3|X_3) \sum_{X_2} P(X_3|X_2)P(X_2|cold_2)
 \end{aligned}$$

First we will solve for the summation:

$$\begin{aligned}
 & \sum_{X_2} P(X_3|X_2)P(X_2|cold_2) \\
 &= [( < 0.2, 0.8, 0, 0, 0, 0 > \times 0) + ( < 0, 0.2, 0.8, 0, 0, 0 > \times 1) \\
 & \quad + ( < 0, 0, 0.2, 0.8, 0, 0 > \times 0) + \dots] \\
 &= < 0, 0.2, 0.8, 0, 0, 0 >
 \end{aligned}$$

Bring our answer back into the original equation:

$$\begin{aligned}
 & P(X_3|hot_1, cold_2, cold_3) \\
 &= \alpha P(cold_3|X_3) < 0, 0.2, 0.8, 0, 0, 0 > \\
 &= \alpha < 0, 1, 1, 0, 1, 1 > \times < 0, 0.2, 0.8, 0, 0, 0 > \\
 &= \alpha < 0, 0.2, 0.8, 0, 0, 0 > \\
 &= < 0, 0.2, 0.8, 0, 0, 0 >
 \end{aligned}$$

Therefore, on the third day, we find that the probability of the rover being on B is 0.2, and the probability that it is on C is 0.8. Any other position has a probability of 0.

$$\text{b) } P(X_2|hot_1, cold_2, cold_3)$$

$$= \alpha P(X_2|hot_1, cold_2)P(cold_3|X_2)$$

$$= \alpha < 0, 1, 0, 0, 0, 0 > \times P(cold_3|X_2)$$

Solving for  $P(cold_3|X_2)$ :

$$= \sum_{X_3} P(cold_3|x_3)P(|x_3)P(x_3|X_2)$$

$$= (0 * 1 * < 0.2, 0, 0, 0, 0, 0 >) +$$

$$(1 * 1 * < 0.8, 0.2, 0, 0, 0, 0 >) +$$

$$(1 * 1 * < 0, 0.8, 0.2, 0, 0, 0, 0 >) +$$

$$(0 * 1 * < 0, 0, 0.8, 0.2, 0, 0, 0 >) +$$

$$(1 * 1 * < 0, 0, 0, 0.8, 0.2, 0 >) +$$

$$(1 * 1 * < 0, 0, 0, 0, 0.8, 0.2 >)$$

$$= < 0.8, 1, 0.2, 0.8, 1, 0.2 >$$

Back to the previous equation:

$$\alpha < 0, 1, 0, 0, 0, 0 > \times P(cold_3|X_2)$$

$$= \alpha < 0, 1, 0, 0, 0, 0 > \times < 0.8, 1, 0.2, 0.8, 1, 0.2 >$$

$$= \alpha < 0, 1, 0, 0, 0, 0 >$$

$$= < 0, 1, 0, 0, 0, 0 >$$

Thus,

$$P(X_2|hot_1, cold_2, cold_3) = < 0, 1, 0, 0, 0, 0 >$$

c) Finding MLE given ( $E_1 = hot, E_2 = cold, E_3 = cold$ )

Our first point of explanation is given because we know the rover fell into A on the first day:

$$V_1(X) = P(X_1) = \langle 1, 0, 0, 0, 0, 0 \rangle$$

Next solving for day 2:

$$\begin{aligned} V_2(X) &= \max_1 [V_1(x_1)P(X_2|x_1)P(E_2 = cold|X_2)] \\ &= \max[1 * 0.2 * 0, 1 * 0.8 * 1, 1 * 0 * 1, \dots] \\ &= \max[\langle 0, 0.8, 0, 0, 0, 0 \rangle] \\ &= 0.8 \end{aligned}$$

That being that our most likely explanation for day 2 is that the rover went to B. Then, for day 3:

$$\begin{aligned} V_3(X) &= \max_2 [V_2(x_2)P(X_3|x_2)P(E_3 = cold|X_3)] \\ &= \max[0, 0.8 * 0.2 * 1, 0.8 * 0.8 * 1, 0 \dots] \\ &= \max[\langle 0, 0.16, 0.64, 0, 0, 0 \rangle] \\ &= 0.64 \end{aligned}$$

From this we can gather that the rover most likely went to C on day 3. Therefore, our most likely explanation is:

Day 1: *A*

Day 2: *B*

Day 3: *C*

d)  $P(hot_4, hot_5, cold_6 | hot_1, cold_2, cold_3)$

Given from previous information on the distribution of the rover to day 3, we know the rover must be on either B or C. Knowing this we can create paths that the rover could have taken.

Starting Point Day 3	Path	Probability
<i>B</i>	<i>BBB</i>	$\alpha 0.008$
<i>B</i>	<i>BBC</i>	$\alpha 0.032$
<i>B</i>	<i>BCC</i>	$\alpha 0.032$
<i>B</i>	<i>BCD</i>	$\alpha 0.128$
<i>B or C</i>	<i>CCC</i>	$\alpha 0.008$
<i>B or C</i>	<i>CCD</i>	$\alpha 0.032$
<i>B or C</i>	<i>CDD</i>	$\alpha 0.032$
<i>B or C</i>	<i>CDE</i>	$\alpha 0.128$
<i>C</i>	<i>DDD</i>	$\alpha 0.008$
<i>C</i>	<i>DDE</i>	$\alpha 0.032$
<i>C</i>	<i>DEE</i>	$\alpha 0.032$
<i>C</i>	<i>DEF</i>	$\alpha 0.128$

Where  $\alpha = 1.666$

From this table, we can see that the ideal path for  $hot_4, hot_5, cold_6$  given we are cold on day 3 is *DDE* when starting on *C*.

Therefore,

$$P(hot_4, hot_5, cold_6 | hot_1, cold_2, cold_3) = \alpha 0.032 = .0533$$

e) Solve  $P(X_4|hot_1, cold_2, cold_3)$  and  $P(X_5|hot_1, cold_2, cold_3)$

$$\begin{aligned} & P(X_4|hot_1, cold_2, cold_3) \\ &= \sum_{X_3} P(X_3|hot_1, cold_2, cold_3)P(X_4|X_3) \end{aligned}$$

Solving for the summation first:

$$\begin{aligned} & \sum_{X_3} P(X_3|hot_1, cold_2, cold_3) \\ &= (0* < 0.2, 0.8, 0, 0, 0, 0 >) \\ &+ (0.2* < 0, 0.2, 0.8, 0, 0, 0 >) \\ &+ (0.8* < 0, 0, 0.2, 0.8, 0, 0 >) \\ &+ (0* < 0, 0, 0, 0.2, 0.8, 0 >) \\ &+ (0* < 0, 0, 0, 0, 0.2, 0.8 >) \\ &+ (0* < 0, 0, 0, 0, 0, 0.2 >) \\ &= < 0, 0.04, 0.32, 0.64, 0, 0 > \end{aligned}$$

Thus,

$$P(X_4|hot_1, cold_2, cold_3) = < 0, 0.04, 0.32, 0.64, 0, 0 >$$

Next, solving for  $P(X_5|hot_1, cold_2, cold_3)$

$$\begin{aligned} &= \sum_{X_4} P(X_4|hot_1, cold_2, cold_3)P(X_5|X_4) \\ &= (0* < 0.2, 0.8, 0, 0, 0, 0 >) \\ &+ (0.04* < 0, 0.2, 0.8, 0, 0, 0 >) \\ &+ (0.32* < 0, 0, 0.2, 0.8, 0, 0 >) \\ &+ (0.64* < 0, 0, 0, 0.2, 0.8, 0 >) \\ &+ (0* < 0, 0, 0, 0, 0.2, 0.8 >) \\ &+ (0* < 0, 0, 0, 0, 0, 0.2 >) \\ &= < 0, 0.008, 0.096, 0.384, 0.512, 0 > \end{aligned}$$

Finally coming to:

$$P(X_5|hot_1, cold_2, cold_3) = < 0, 0.008, 0.096, 0.384, 0.512, 0 >$$

## Question 5

- a) Finding  $P(X_t|e, \alpha)$  where  $\alpha$  denotes our action. Each matrix denotes the probability that we are at that position for each step. For full work, please check the included documents labeled "question5-partA\_work" under the folder "q5".

- $P(X_1|N, R)$

$$\begin{aligned}
 &= P(N|X_1) \sum_{X_0} P(X_1|X_0)P(X_0) \\
 &= \begin{bmatrix} 0.00131 & 0.01315 & 0.025 \\ 0.02368 & 0.2368 & 0.45 \\ 0.2368 & & 0.013 \end{bmatrix}
 \end{aligned}$$

- $P(X_2|N, N, R)$

$$\begin{aligned}
 &= P(N|X_2) \sum_{X_1} P(X_2|X_1)P(X_1|N) \\
 &= \begin{bmatrix} 7.66 * 10^{-6} & 1.46 * 10^{-4} & 0.0021 \\ 0.0025 & 0.0474 & 0.698 \\ 0.249 & & 7.6 * 10^{-4} \end{bmatrix}
 \end{aligned}$$

- $P(X_3|N, N, H, D)$

$$\begin{aligned}
 &= P(H|X_3) \sum_{X_2} P(X_3|X_2)P(X_2|N, N) \\
 &= \begin{bmatrix} 1.78 * 10^{-6} & 2.247 * 10^{-5} & 1.796 * 10^{-5} \\ 2.198 * 10^{-5} & 0.00406 & 0.00613 \\ 0.02148 & & 0.9682 \end{bmatrix}
 \end{aligned}$$

- $P(X_4|N, N, H, H, D)$

$$\begin{aligned}
 &= P(H|X_4) \sum_{X_3} P(X_4|X_3)P(X_3|N, N, H) \\
 &= \begin{bmatrix} 1.825 * 10^{-7} & 2.304 * 10^{-6} & 1.023 * 10^{-7} \\ 2.165 * 10^{-7} & 2.324 * 10^{-4} & 3.583 * 10^{-5} \\ 0.00122 & & 0.9985 \end{bmatrix}
 \end{aligned}$$

Given this information, we can conclude that we will have ended up in the bottom right of the map (3,3) with a probability of 0.9985.

- b) For example readings generated, please refer to the included python scripts and the files under the folder "q5" and the generated results under "results" within "q5".