CSE 250A HW3

Jiping Lin A15058075

1. (a) when $t=1$, $P(X_2=j|X_1=i)=[A]_{ij}$.

Suppose $P(X_{k+1}=j|X_1=i)=[A^k]_{ij}$ for some $k>1$,

then
$$
\begin{aligned}
P(X_{k+2}=j|X_1=i) &= \sum_{\ell=1}^{m} P(X_{k+2}=j, X_{k+1}=\ell \mid X_1=i) \\
&= \sum_{\ell=1}^{m} P(X_{k+1}=\ell|X_1=i)P(X_{k+2}=j|X_{k+1}=\ell, X_1=i) \quad \text{「Product Rule」} \\
&= \sum_{\ell=1}^{m} [A^k]_{i\ell} P(X_{k+2}=j|X_{k+1}=\ell) \quad (\text{by d.sep (1)) } X_1 \to X_{k+2} \mid X_{k+1} \\
&= \sum_{\ell=1}^{m} [A^k]_{i\ell}[A]_{\ell j} \\
&= [A^{k+1}]_{ij},
\end{aligned}
$$

By induction, we have $P(X_{t+1}=j|X_1=i)=[A^t]_{ij}$, $\forall t\geq 1$.

(b)
```
int prob(int[][] A, int t, int j, int i){
    if (t==1) return A[i][j];
    int res = 0;
    for (int k = 0; k < A.length; k++)          (*)
        res += prob(A, t-1, k, i) * prob(A, 1, j, k);
    return res;
}
```

Since $[A^t]_{ij}=\sum_{k=1}^{m}[A^{t-1}]_{ik}[A]_{kj}$, we use recursion to compute each term in summation, each row * col is $O(m)$. prob(A, t-1, k, i) needs $O(t)$ to reduce to base case, and sum takes m steps, so it's $O(m^2t)$.

(c) Similarly, we split the A^k to $A^{k/2}$ and $A^{k/2}$ in recursive steps:

```
int prob2(int[][] A, int t, int j, int i){
    if (t==1) return A[i][j];
    int res = 0;
    if((t&1)==1){
        for (int k=0; k<A.length; k++){
            res += prob2(A, t-1, k, i) * prob2(A, 1, j, k);
        }
    } else {
        for(int k=0; k<A.length; k++){
            res += prob2(A, t/2, k, i)*prob2(A, t/2, j, k);
        }
    }
    return res;
}
```

This is $O(m^3 \log t)$ since recursion takes $\log(t)$ but in summation we don't always have base case ($t=1$), so it is $O(m^2)$

(d) Instead of going over the entire row, we only look at nonzero entries.
We add following to function in (b):

```
Set<Integer> nonzero = new HashSet<>();
for (int k=0; k<A.length; k++)
    if (A[i][k] != 0) nonzero.add(k);
```

then change (*) to `for (int k : nonzero)`.
This is $O(smt)$ since the row * col becomes $O(s)$.

(e) $P(X_1=i \mid X_{t+1}=j) = \dfrac{P(X_{t+1}=j \mid X_1=i)\, P(X_1=i)}{P(X_{t+1}=j)}$, ⌜Baye's⌟

where
$$P(X_{t+1}=j) = \sum_{i=1}^{m} P(X_{t+1}=j, X_1=i)$$
$$= \sum_{i=1}^{m} P(X_{t+1}=j \mid X_1=i)\, P(X_1=i)$$
$$= \sum_{i=1}^{m} [A^t]_{ij}\, P(X_1=i),$$

So
$$P(X_1=i \mid X_{t+1}=j) = \frac{[A^t]_{ij}\, P(X_1=i)}{\sum_{k=1}^{m} [A^t]_{kj}\, P(X_1=k)}$$

2. (a)
$$P(Y_1 \mid X_1) = \sum_x P(Y_1, X_0=x \mid X_1)$$
$$= \sum_x P(Y_1 \mid X_0=x, X_1)\, P(X_0=x \mid X_1)$$
$$= \sum_x P(Y_1 \mid X_0=x, X_1)\, P(X_0=x)$$

(b)
$$P(Y_1) = \sum_x P(Y_1, X_1=x)$$
$$= \sum_x P(Y_1 \mid X_1=x)\, P(X_1=x)$$

(c) $P(X_n \mid Y_1, \cdots, Y_{n-1}) = P(X_n)$

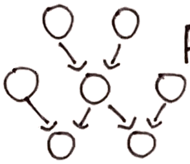
By d.sep(3), $X_n \to Y_i \mid \emptyset$ is blocked by Y_n, $\forall i$.

(d)
$$P(Y_n \mid X_n, Y_1, \cdots, Y_{n-1}) = \sum_x P(Y_n, X_{n-1}=x \mid X_n, Y_1, \cdots, Y_{n-1})$$
$$= \sum_x P(Y_n \mid X_n, Y_1, \cdots, Y_{n-1}, X_{n-1}=x)\, P(X_{n-1}=x \mid X_n, Y_1, \cdots, Y_{n-1})$$
$$= \sum_x P(Y_n \mid X_n, X_{n-1})\, P(X_{n-1}=x \mid Y_1, \cdots, Y_{n-1})$$

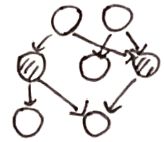
⌜$Y_n \to Y_1, \cdots, Y_{n-1} \mid X_n, X_{n-1}$ by d.sep(1), $X_{n-1} \to X_n \mid Y_1, \cdots, Y_{n-1}$ by d.sep(3)⌟

(e)
$$P(Y_n \mid Y_1, \cdots, Y_{n-1}) = \sum_x P(Y_n, X_n=x \mid Y_1, \cdots, Y_{n-1})$$
$$= \sum_x P(X_n=x \mid Y_1, \cdots, Y_{n-1})\, P(Y_n \mid X_n=x, Y_1, \cdots, Y_{n-1})$$
$$= \sum_x \underbrace{P(X_n=x)}_{\text{by (c)}}\, \underbrace{P(Y_n \mid X_n=x, Y_1, \cdots, Y_{n-1})}_{\text{known from (d)}}$$

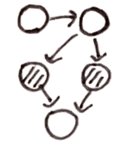
3.



Polytree



X Polytree

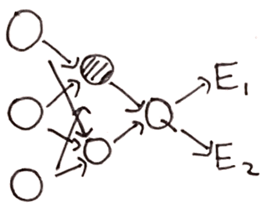


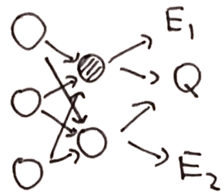
x Polytree

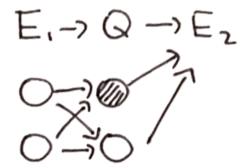
Polytree

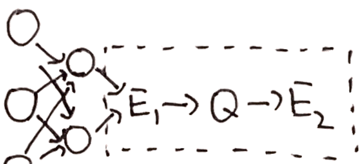
Polytree.

4.

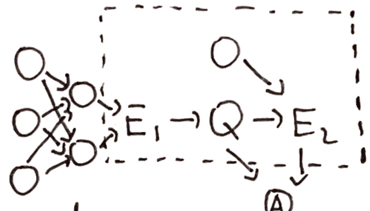








d.sep(1) by E_1



d.sep (1) by E_1.

5.

Y	P(Y|X=0)	P(Y|X=1)	P(Z₁=1|Y)	P(Z₂=1|Y)
1	(0.25)(0.5)(0.75) = 0.09375	(0.5)(0.75)(0.25) = 0.09375	0.9	0.1
2	(0.75)(0.5)(0.75) = 0.28125	(0.5)(0.75)(0.25) = 0.09375	0.8	0.2
3	(0.25)(0.5)(0.75) = 0.09375	(0.5)(0.25)(0.25) = 0.03125	0.7	0.3
4	(0.25)(0.5)(0.25) = 0.03125	(0.5)(0.75)(0.75) = 0.28125	0.6	0.4
5	(0.75)(0.5)(0.75) = 0.28125	(0.5)(0.25)(0.25) = 0.03125	0.5	0.5
6	(0.75)(0.5)(0.25) = 0.09375	(0.5)(0.75)(0.75) = 0.28125	0.4	0.6
7	(0.25)(0.5)(0.25) = 0.03125	(0.5)(0.25)(0.75) = 0.09375	0.3	0.7
8	(0.75)(0.5)(0.25) = 0.09375	(0.5)(0.25)(0.75) = 0.09375	0.2	0.8

6. (a)
$$\sum_z P(Z=z|B_1,\cdots,B_n) = \sum_z \left(\frac{1-\alpha}{1+\alpha}\right)\alpha^{|z-f(B)|}$$
$$= \left(\frac{1-\alpha}{1+\alpha}\right)\sum_z \alpha^{|z-f(B)|}$$
$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left[\sum_{z\le f(B)} \alpha^{f(B)-z} + \sum_{z\ge f(B)} \alpha^{z-f(B)} - \alpha^{f(B)-f(B)}\right]$$
$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left[\sum_{z\le f(B)} \frac{\alpha^{f(B)}}{\alpha^{z}} + \sum_{z\ge f(B)} \frac{\alpha^{z}}{\alpha^{f(B)}} - 1\right]$$
$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left[\frac{1}{1-\alpha} + \frac{1}{1-\alpha} - 1\right]$$
$$= \left(\frac{1-\alpha}{1+\alpha}\right)\left(\frac{1+\alpha}{1-\alpha}\right) = 1$$

(b)-(d) In code part

7. (a) $P(B|A,C,D) = \dfrac{P(D|B,A,C)\,P(B|A,C)}{P(D|A,C)}$ ⌈Swap B, D Baye's Rule⌋

where

$P(D|B,A,C) = P(D|B,C)$ ⌈ $D \to A|B,C$, d.sep (1) ⌋

$P(B|A,C) = P(B|A)$ ⌈ $B \to C|A$, d.sep(3) ⌋

$P(D|A,C) = \sum_b P(D,B=b|A,C) = \sum_b P(D|B=b,A,C)P(B=b|A,C) = \sum_b P(D|B=b,C)P(B=b|A)$

So $P(B|A,C,D)$
$$= \frac{P(D|B,C)\,P(B|A)}{\sum_b P(D|B=b,C)P(B=b|A)}$$

(b) $P(B|A,C,D,E,F) = P(B|A,C,D)$ 「$B \to E,F | A\,C\,D$ by d.sep(2),
$B \to E$ blocked by C,
$B \to F$ blocked by A.」

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

$$= \underbrace{P(B|A,C,D)}_{\text{by (a)}}\ \underbrace{P(E|C)}_{\substack{E\to A,F,D|C \\ \text{by d.sep(2)}}}\ \underbrace{P(F|A)}_{\substack{F\to C,D|A \\ \text{by d.sep(2).}}}$$

8. (a)
$$P(Q=q|E=e) \approx \frac{\sum_{t=1}^{T} I(q,q_t)\,P(E=e|Y=y_t, Z=z_t)}{\sum_{t=1}^{T} P(E=e|Y=y_t, Z=z_t)}$$

where y_t is from $P(Y|X=x_t, Q=q_t)$, $x_t \sim P(X)$,
z_t is from $P(Z|Q=q_t)$ $\quad q_t \sim P(Q)$.

(b)
$$P(Q_1=q_1, Q_2=q_2) \approx \frac{\sum_{t=1}^{T} I(q_1,q_{1t})\,I(q_2,q_{2t})\,P(E_1=e_1|Q_1=q_{1t}, X=x_t)\,P(E_2=e_2|E_1=e_1, Z=z_t)}{\sum_{t=1}^{T} P(E_1=e_1|Q_1=q_{1t}, X=x_t)\,P(E_2=e_2|E_1=e_1, Z=z_t)}$$

where q_{1t} is from $P(Q_1)$, x_t is from $P(X)$, y_t is from $P(Y)$
z_t is from $P(Z|X=x_t, Y=y_t)$