PGM HWK #2

Haekyu Park

March 5, 2020.

Problem 1 (Variable Elimination)

We can factorize $P(w,o,c,m,f,g,b,s,h)$ as

$P(w)\,P(o)\,P(c)\,P(m)\,P(f|w,o)\,P(g|c,m)\,P(b|f,g)\,P(s|f,b)\,P(h|s,g)$.

1. $P(S=T \mid C=T)$. $\longrightarrow$ The answer is $\boxed{0.782978}$ (see p4).

$$P(S=T, C=T)$$

$$= \sum_{\substack{w,o,m,\\ f,g,b,h}} P(w)\,P(o)\,0.4\,P(m)\,P(f|w,o)\,P(g|C=T,m)\,P(b|f,g)\,P(S=T|f,b) \times P(h|S=T,g).$$

① Eliminate H.

$$m_h(g) = \sum_h P(h \mid S=T, g) = 1.$$

$\downarrow$

$$P(S=T, C=T) = 0.4 \times \sum_{\substack{w,o,m,\\ f,g,b}} P(w)\,P(o)\,P(m)\,P(f|w,o)\,P(g|C=T,m)\,P(b|f,g) \times P(S=T|f,b).$$

② Eliminate M.

$$m_m(g) = \sum_m P(m) \cdot P(g \mid C=T, m)$$

- $$m_m(g=T) = \sum_m P(m)\, P(g=T \mid C=T, m)$$
$$= (0.1 \cdot 0.1) + (0.9 \cdot 0.9) = 0.82$$

- $$m_m(g=F) = \sum_m P(m)\, P(g=F \mid C=T, m)$$
$$= (0.1 \cdot 0.9) + (0.9 \cdot 0.1) = 0.18.$$

↓

$$P(S=T, C=T) = 0.4 \sum_{\substack{w,o,\\ f,g,b}} P(w)\, P(o)\, P(f \mid w,o)\, P(b \mid f,g)\, P(S=T \mid f,b) \times m_m(g).$$

③ Eliminate w, o

$$m_{wo}(f) = \sum_{w,o} P(w) P(o)\, P(f \mid w,o)$$

- $$m_{wo}(f=T) = (0.3 \cdot 0.6 \cdot \underline{0.9}) + (0.3 \cdot 0.4 \cdot \underline{0.6})$$
$$+ (0.7 \cdot 0.6 \cdot \underline{0.8}) + (0.7 \cdot 0.4 \cdot \underline{0.5}) = \underline{0.71}$$

- $$m_{wo}(f=F) = (0.3 \cdot 0.6 \cdot 0.1) + (0.3 \cdot 0.4 \cdot 0.4)$$
$$+ (0.7 \cdot 0.6 \cdot 0.2) + (0.7 \cdot 0.4 \cdot 0.5) = 0.29.$$

↓

$$P(S=T, C=T) = 0.4 \sum_{f,g,b} P(b \mid f,g)\, P(S=T \mid f,b)\, m_{wo}(f)\, m_m(g)$$

④ Eliminate b

$$m_b(f,g) = \sum_b P(b \mid f,g)\, P(S=T \mid f,b).$$

- $m_b(f=T, g=T) = \sum_b P(b \mid f=T, g=T)\, P(S=T \mid f=T, b)$

$$= (0.9 \cdot 0.95) + (0.1 \cdot 0.6) = 0.915$$

- $m_b(f=T, g=F) = \sum_b P(b \mid f=T, g=F)\, P(S=T \mid f=T, b)$

$$= (0.7 \cdot 0.95) + (0.3 \cdot 0.6) = 0.845$$

- $m_b(f=F, g=T) = \sum_b P(b \mid f=F, g=T)\, P(S=T \mid f=F, b)$

$$= (0.75 \cdot 0.7) + (0.25 \cdot 0.1) = 0.55$$

- $m_b(f=F, g=F) = \sum_b P(b \mid f=F, g=F)\, P(S=T \mid f=F, b)$

$$= (0.2 \cdot 0.7) + (0.8 \cdot 0.1) = 0.22.$$

$\downarrow$

$$P(S=T, C=T) = 0.4 \sum_{f,g} m_b(f,g)\, m_{wo}(f)\, m_m(g)$$

⑤ Eliminate f

$$m_f(g) = \sum_f m_b(f,g) \cdot m_{wo}(f).$$

- $$m_f(g=T) = \sum_f m_b(f, g=T)\, m_{wo}(f)$$

$$= (0.915 \cdot 0.71) + (0.55 \cdot 0.29) = 0.80915$$

- $$m_f(g=F) = \sum_f m_b(f, g=F)\, m_{wo}(f)$$

$$= (0.845 \cdot 0.71) + (0.22 \cdot 0.29) = 0.66375.$$

$\downarrow$

$$P(S=T, C=T) = 0.4 \sum_g m_f(g)\, m_m(g)$$

⑥ For all g,

$$P(S=T, C=T) = 0.4 \sum_g m_f(g)\, m_m(g)$$

$$= 0.4 \cdot \{(0.80915 \cdot 0.82) + (0.66375 \cdot 0.18)\}$$

$$= 0.3131912$$

$$\therefore P(S=T \mid C=T) = \frac{P(S=T, C=T)}{P(C=T)} = \frac{0.313912}{0.4}$$

$$= \boxed{0.782978}$$

$\boxed{\text{* Ordering}}$: H, M, $\underbrace{\text{W, O}}_{\text{at the same time}}$, b, f, g

$\boxed{\text{* factors at each step}}$: see how P(S=T, C=T) be updated for each ①, ②, ⋯, ⑥.

2. $P(F=T \mid G=T)$

$$P(F=T \mid G=T) = P(F=T) = \boxed{0.71}$$

F and G are independent

See 1-⑪.

3. $P(M=T \mid G=T)$

$$P(M=T \mid G=T) = \frac{P(M=T, G=T)}{P(G=T)} = \frac{①}{②}$$

① $P(M=T, G=T) = P(G=T \mid M=T)\, P(M=T)$

$$= P(M=T) \cdot \sum_c P(G=T \mid M=T, c)\, P(c)$$

$$= 0.1 \times (0.4 \cdot 0.1 + 0.6 \cdot 0.01) = 0.0046$$

② $P(G=T) = \sum_{c,m} P(G=T \mid c,m)\, P(c)\, P(m)$

$$= (0.4 \cdot 0.1 \cdot 0.1) + (0.4 \cdot 0.9 \cdot 0.9)$$

$$+ (0.6 \cdot 0.1 \cdot 0.01) + (0.6 \cdot 0.9 \cdot 0.7)$$

$$= 0.7066$$

$$\therefore P(M=T \mid G=T) = \frac{①}{②} = \frac{0.0046}{0.7066} = \boxed{0.00651}$$

4. $P(M=T \mid G=T, S=T)$

$$P(M=T \mid G=T, S=T)$$

$$= P(M=T \mid G=T) \quad \text{) } M \perp S \mid G$$

$$= \frac{P(M=T, G=T)}{P(G=T)}$$

$$= \frac{\overbrace{P(G=T \mid M=T)}^{①} \; \overbrace{P(M=T)}^{0.1}}{\underbrace{P(G=T)}_{②}}$$

$$① = P(G=T \mid M=T) = \sum_{c} P(G=T \mid M=T, c) \cdot P(c).$$

$$= (0.4 \cdot 0.1) + (0.6 \cdot 0.01) = 0.046.$$

$$② = P(G=T) = \sum_{c,m} P(G=T \mid c, m)\, P(c)\, P(m)$$

$$= 0.7066. \quad \leftarrow \text{See } 3\text{-}②.$$

$$\therefore P(M=T \mid G=T, S=T) = \frac{① \times 0.1}{②} = \frac{0.046 \times 0.1}{0.7066}$$

$$= \boxed{0.00651}$$

5. $P(W=T \mid G=T, B=F, S=T)$

$P(W=T, G=T, B=F, S=T)$

$$= \sum_{o,c,m,f,h} P(w=T, o, c, m, f, g=T, b=F, s=T, h).$$

$$= \sum_{o,c,m,f,h} \underline{P(w=T)} P(o) P(c) P(m) \underline{P(f \mid w=T, o)} P(g=T \mid c, m) \times P(b=F \mid f, g=T) \cdot P(S=T \mid f, b=F) \times P(h \mid S=T, G=T).$$

① Eliminate ω.

$$m_\omega(o, f) = \sum_\omega P(\omega=T) P(f \mid \omega=T, o)$$

- $m_\omega(o=T, f=T) = 0.3 \cdot 0.9 = 0.27$
- $m_\omega(o=T, f=F) = 0.3 \cdot 0.1 = 0.03$
- $m_\omega(o=F, f=T) = 0.3 \cdot 0.6 = 0.18$
- $m_\omega(o=F, f=F) = 0.3 \cdot 0.4 = 0.12.$

$\rightarrow P(W=T, G=T, B=F, S=T)$

$$= \sum_{o,c,m,f,h} P(o) P(c) P(m) P(g=T \mid c, m) P(b=F \mid f, g=T) \times P(S=T \mid f, b=F) \cdot P(h \mid S=T, G=T) \cdot m_\omega(o, f)$$

Scanned with CamScanner

② Eliminate H.

$$\sum_h P(h \mid S=T, G=T) = 1$$

$\rightarrow P(W=T, G=T, B=F, S=T)$

$$= \sum_{\substack{o,c,\\ m,f}} P(o)P(c)P(m)P(g=T \mid c,m)P(b=F \mid f, g=T) \times P(S=T \mid f, b=F) \cdot m_w(o,f)$$

③ Eliminate F, O

$$\sum_{f,o} P(b=F \mid f, g=T) \cdot P(S=T \mid f, b=F) \cdot m_w(o,f)$$

$$= \sum_f \left(P(b=F \mid f, g=T) P(S=T \mid f, b=F) \sum_o m_w(o,f) \right)$$

$$= \{0.1 \cdot 0.6 \cdot (0.27 + 0.18)\} + \{0.25 \cdot 0.7 \cdot (0.03 + 0.12)\}$$

$$= 0.05325$$

$\rightarrow P(W=T, G=T, B=F, S=T)$

$$= 0.05325 \sum_{m,c} P(c) \cdot P(m) \cdot P(g=T \mid c, m)$$

$$= 0.05325 \cdot \{(0.4 \cdot 0.1 \cdot 0.1) + (0.4 \cdot 0.9 \cdot 0.9) + (0.6 \cdot 0.1 \cdot 0.01) + (0.6 \cdot 0.9 \cdot 0.7)\}$$

$$= \boxed{0.03762645}$$

Problem 2

The matrix is

[[7 4 10 15 21 10 10]
 [10 6 7 7 16 10 12]
 [6 6 10 16 11 6 15]
 [13 17 12 18 19 12 18]
 [11 7 15 12 6 9 14]]

as printed out as follows.

```
(base)  20-03-07  19:22  ~/GoogleDrive/Gatech/pgm/HW/hw2/hw2-codes  python problem2.py
problem2.py:187: RuntimeWarning: divide by zero encountered in log
  omega[0, :] = np.log(phi) + np.log(E[:, V[0]])
problem2.py:195: RuntimeWarning: divide by zero encountered in log
  probability = omega[char_th - 1] + np.log(T[:, state]) + np.log(E[state, V[char_th]])
[[ 7  4 10 15 21 10 10]
 [10  6  7  7 16 10 12]
 [ 6  6 10 16 11  6 15]
 [13 17 12 18 19 12 18]
 [11  7 15 12  6  9 14]]
```

Problem 3.

Let's denote $\gamma_c(b) := \max_c p(b|c)\, p(c)$.

① If $b =$ true

$$\gamma_c(b=t) = \max_c p(b=\text{true} \mid c)\, p(c)$$

$$= \max \{ p(b=t \mid c=t)\, p(c=t),\ p(b=t \mid c=f)\, p(c=f) \}$$

$$= \max \{ 0.75 \times 0.4,\ 0.1 \times 0.6 \}$$

$$= \max \{ 0.3,\ 0.06 \} = 0.3$$

② If $b =$ false

$$\gamma_c(b=f) = \max_c p(b=f \mid c)\, p(c)$$

$$= \max \{ p(b=f \mid c=t)\, p(c=t),\ p(b=f \mid c=f)\, p(c=f) \}$$

$$= \max \{ (1-0.75)\, 0.4,\ (1-0.1) \times 0.6 \}$$

$$= \max \{ 0.1,\ 0.54 \} = 0.54$$

The optimal state is; $b^* =$ false

Let's denote $\gamma_b(a) = \max_b p(a|b)\, \gamma_c(b)$

① If $a = \text{true}$

$$\gamma_b(a=t) = \max_b p(a=t\,|\,b)\, \gamma_c(b)$$

$$= \max \{ p(a=t\,|\,b=t)\, \gamma_c(b=t),$$

$$p(a=t\,|\,b=f)\, \gamma_c(b=f) \}$$

$$= \max \{ 0.3 \times 0.3,\ 0.2 \times 0.54 \} = \underbrace{0.108}$$

② If $a = \text{false}$

$$\gamma_b(a=f) = \max_b p(a=f\,|\,b)\, \gamma_c(b)$$

$$= \max \{ p(a=f\,|\,b=t)\, \gamma_c(b=t),$$

$$p(a=f\,|\,b=f)\, \gamma_c(b=f) \}$$

$$= \max \{ \overset{b=t}{(1-0.3) \times 0.3},\ \underbrace{\overset{b=f}{(1-0.2) \times 0.54}} \} = \underbrace{0.432}$$

The optimal state is ; $a^* = \text{false}$

$b^* = \text{false}$

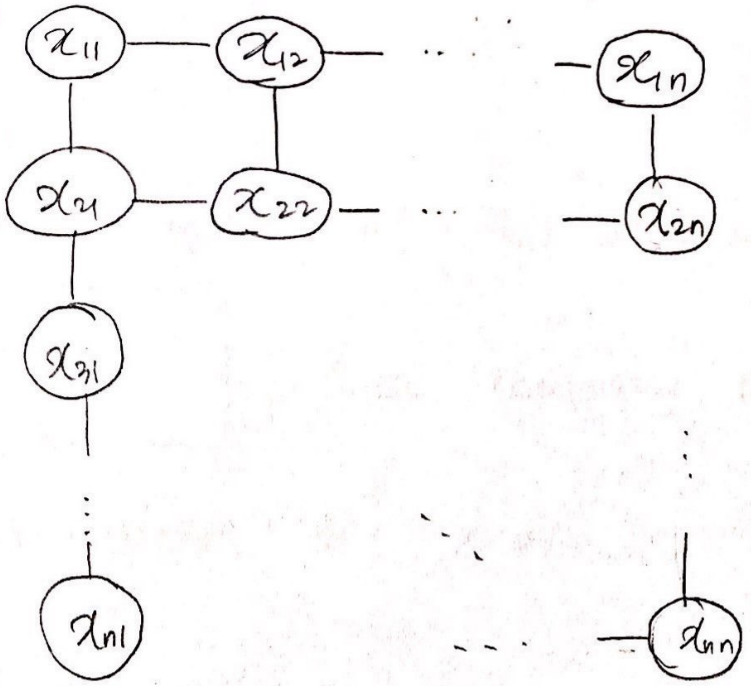
$c^* = \arg\max_c \gamma_c(b = b^* = f) = \text{false}.$

$$\boxed{\therefore\ a^* = \text{false},\ b^* = \text{false},\ c^* = \text{false}}$$

Scanned with CamScanner

Problem 4.

The Ising graph would look like as follows.



indicating column

The node potential for each X_c is :

$$\phi(X_c) = \prod_{r=1}^{n-1} e^{\mathbb{I}[x_{rc},\, x_{r+1,c}]}$$

indicating row

The edge potential for each edge is :

$$\phi(X_i, X_j) = \prod_{j=1}^{n} e^{\mathbb{I}[x_{ji},\, x_{j,i+1}]}$$

The message passing would be as follows.

$$m_{i,i+1} = \sum_i \phi(X_i)\, \phi(X_i, X_j)\, m_{i-1,i}$$

The time complexity of the message passing is

$\boxed{O(n \cdot 2^{2n})}$, since computing $m_{i,i+1}$ takes $O(2^{2n})$

and we repeat this n times.

When $n = 10$, $\log(Z)$ is $\boxed{183,245}$.

```
(base)  20-03-07  19:24  ~/GoogleDrive/Gatech/pgm/HW/hw2/hw2-codes  python problem4.py
183.24517348338284
```

Problem 5

1. The marginals of the symptoms are:
   p(s_1=1) = 0.441834
   p(s_2=1) = 0.456675
   p(s_3=1) = 0.441405
   p(s_4=1) = 0.491275
   p(s_5=1) = 0.493892
   p(s_6=1) = 0.657483
   p(s_7=1) = 0.504563
   p(s_8=1) = 0.268693
   p(s_9=1) = 0.649077
   p(s_10=1) = 0.49074
   p(s_11=1) = 0.422552
   p(s_12=1) = 0.429096
   p(s_13=1) = 0.545021
   p(s_14=1) = 0.632959
   p(s_15=1) = 0.42954
   p(s_16=1) = 0.458794
   p(s_17=1) = 0.427559
   p(s_18=1) = 0.404255
   p(s_19=1) = 0.582093
   p(s_20=1) = 0.589591
   p(s_21=1) = 0.76127
   p(s_22=1) = 0.695588
   p(s_23=1) = 0.508702
   p(s_24=1) = 0.419962
   p(s_25=1) = 0.351942
   p(s_26=1) = 0.389611
   p(s_27=1) = 0.325973
   p(s_28=1) = 0.469624
   p(s_29=1) = 0.522868
   p(s_30=1) = 0.717312
   p(s_31=1) = 0.524199
   p(s_32=1) = 0.353704
   p(s_33=1) = 0.512679
   p(s_34=1) = 0.529404
   p(s_35=1) = 0.385751
   p(s_36=1) = 0.489095
   p(s_37=1) = 0.633595
   p(s_38=1) = 0.589604
   p(s_39=1) = 0.423164
   p(s_40=1) = 0.528234

2. The marginals $p(s_i = 1)$ can be computed more efficiently, by summation based on the symptom s_i 's parents. The following equation indicates the efficient process, where P_1, P_2, P_3 are the symptoms' parents.

$$p(s_i = 1) = \sum_{P_1, P_2, P_3} p(s_i = 1 | P_1, P_2, P_3) p(P_1) p(P_2) p(P_3)$$

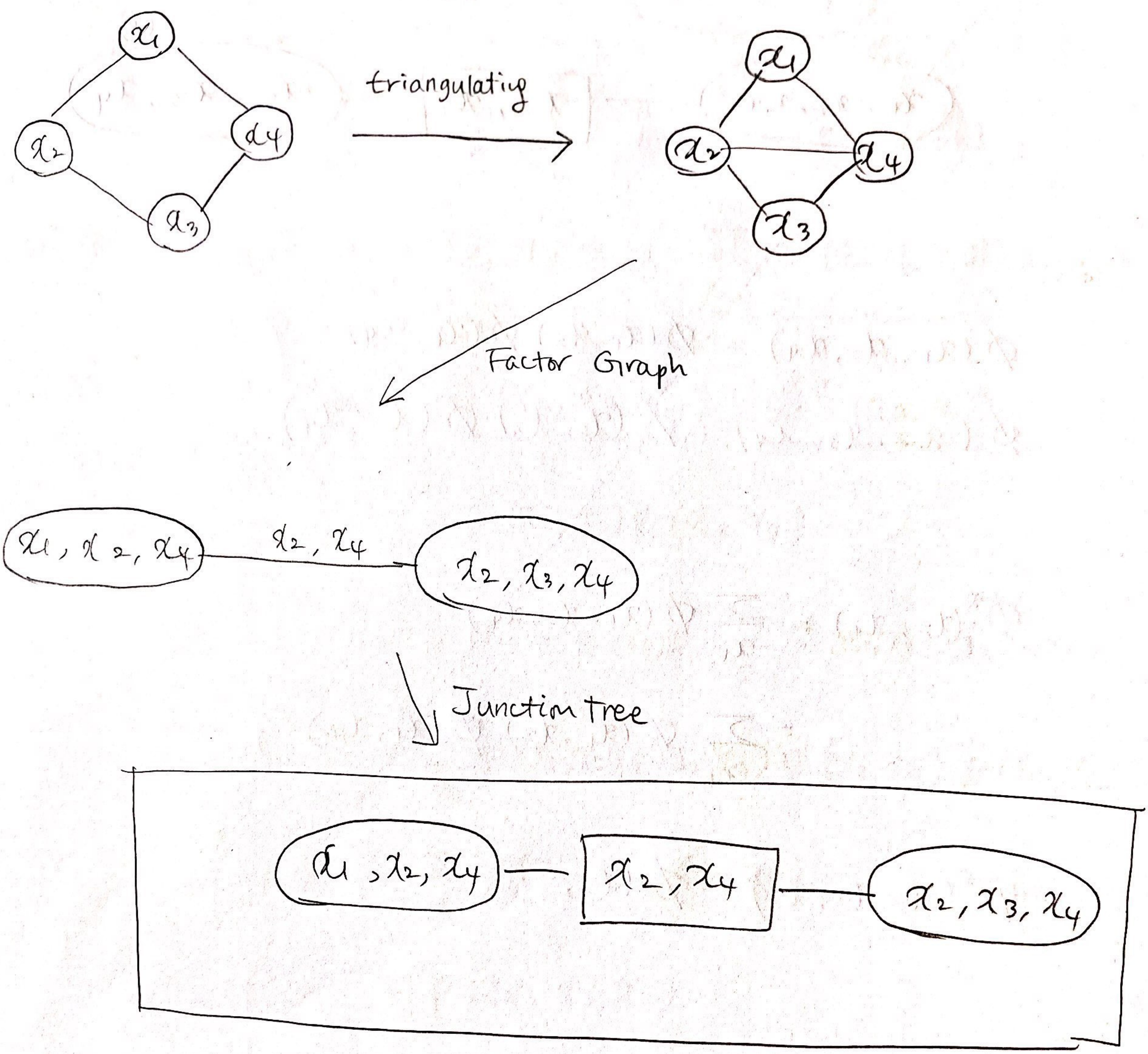
   In the junction tree formalism, it takes 0.25614 seconds. In the efficient method, it takes 0.057125 seconds. The efficient method is faster.

3. The marginals of the diseases given symptoms are:

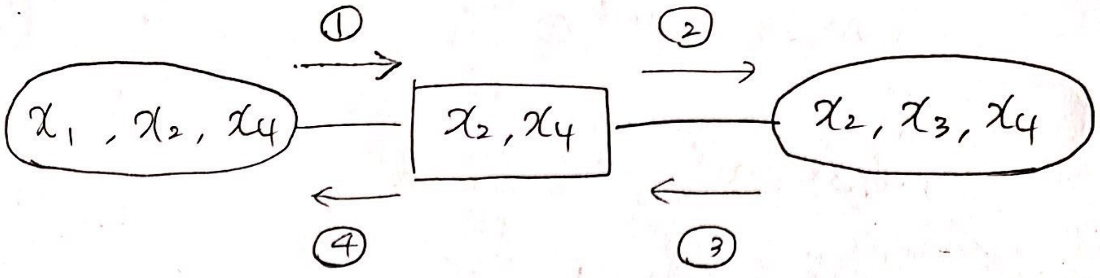
   p(d_1 = 1|s_1:10)=0.0297756
   p(d_2 = 1|s_1:10)=0.38176
   p(d_3 = 1|s_1:10)=0.954235
   p(d_4 = 1|s_1:10)=0.396644
   p(d_5 = 1|s_1:10)=0.496467
   p(d_6 = 1|s_1:10)=0.435154
   p(d_7 = 1|s_1:10)=0.187487
   p(d_8 = 1|s_1:10)=0.701183
   p(d_9 = 1|s_1:10)=0.0431266
   p(d_10 = 1|s_1:10)=0.610313
   p(d_11 = 1|s_1:10)=0.287322
   p(d_12 = 1|s_1:10)=0.489833
   p(d_13 = 1|s_1:10)=0.8996
   p(d_14 = 1|s_1:10)=0.619565
   p(d_15 = 1|s_1:10)=0.920476
   p(d_16 = 1|s_1:10)=0.706096
   p(d_17 = 1|s_1:10)=0.201247
   p(d_18 = 1|s_1:10)=0.908494
   p(d_19 = 1|s_1:10)=0.864967
   p(d_20 = 1|s_1:10)=0.883929

Problem 6

1. A junction tree.



2.



① $\phi^*(x_2, x_4) = \sum_{x_1} \phi(x_1, x_2, x_4) = \sum_{x_1} \phi(x_1, x_2)\phi(x_1, x_4)$

② $\phi^*(x_2, x_3, x_4) = \underbrace{\phi(x_2, x_3, x_4)}_{=\,\phi(x_2, x_3)\phi(x_3, x_4)} \cdot \phi^*(x_2, x_4)$

$$= \phi(x_2, x_3)\phi(x_3, x_4) \sum_{x_1} \phi(x_1, x_2)\phi(x_1, x_4)$$

$$= \sum_{x_1} \phi(x_1, x_2)\phi(x_2, x_3)\phi(x_3, x_4)\phi(x_4, x_1)$$

$$= \sum_{x_1} p(x_1, x_2, x_3, x_4)$$

$$= p(x_2, x_3, x_4)$$

③. $\phi^{**}(x_2, x_4) = \sum_{x_3} \phi^{*}(x_2, x_3, x_4)$ by ②

$= \sum_{x_3} p(x_2, x_3, x_4)$

$= p(x_2, x_4)$

④ $$\phi^{*}(x_1, x_2, x_4) = \frac{\phi(x_1, x_2, x_4)\, \overbrace{\phi^{**}(x_2, x_4)}^{A}}{\underbrace{\phi^{*}(x_2, x_4)}_{B}}$$

$$A = \sum_{x_3} \phi(x_2, x_3)\phi(x_3, x_4) \sum_{x_1} \phi(x_1, x_2)\phi(x_1, x_4)$$

$$B = \sum_{x_1} \phi(x_1, x_2)\phi(x_1, x_4)$$

$$\therefore \phi^{*}(x_1, x_2, x_4) = \frac{\phi(x_1, x_2, x_4)\, A}{B}$$

$$= \phi(x_1, x_2, x_4) \sum_{x_3} \phi(x_2, x_3)\, \phi(x_3, x_4)$$

$$= \sum_{x_3} \phi(x_1, x_2)\, \phi(x_2, x_3)\, \phi(x_3, x_4)\, \phi(x_4, x_1)$$

$$= \sum_{x_3} p(x_1, x_2, x_3, x_4) = p(x_1, x_2, x_4)$$

$$\boxed{\therefore p(x_1) = \sum_{x_2, x_4} \phi^{*}(x_1, x_2, x_4) = \sum_{x_2, x_4} p(x_1, x_2, x_4) = p(x_1)}$$

Scanned with CamScanner

Problem 7.

1.

The junction tree is the clique of all variables $x_1, \cdots, x_T$, since we moralize the graph and connect all variables.

The time complexity of the junction tree algorithm is exponential to the size of the largest cluster in the junction tree, which is the clique itself in this case. Thus, the time complexity is $O(2^T)$

To sum up,

the junction tree is the clique of all variables

and the time complexity of computing $P(x_T)$ by the junction tree algorithm is $\boxed{O(2^T)}$.

2.

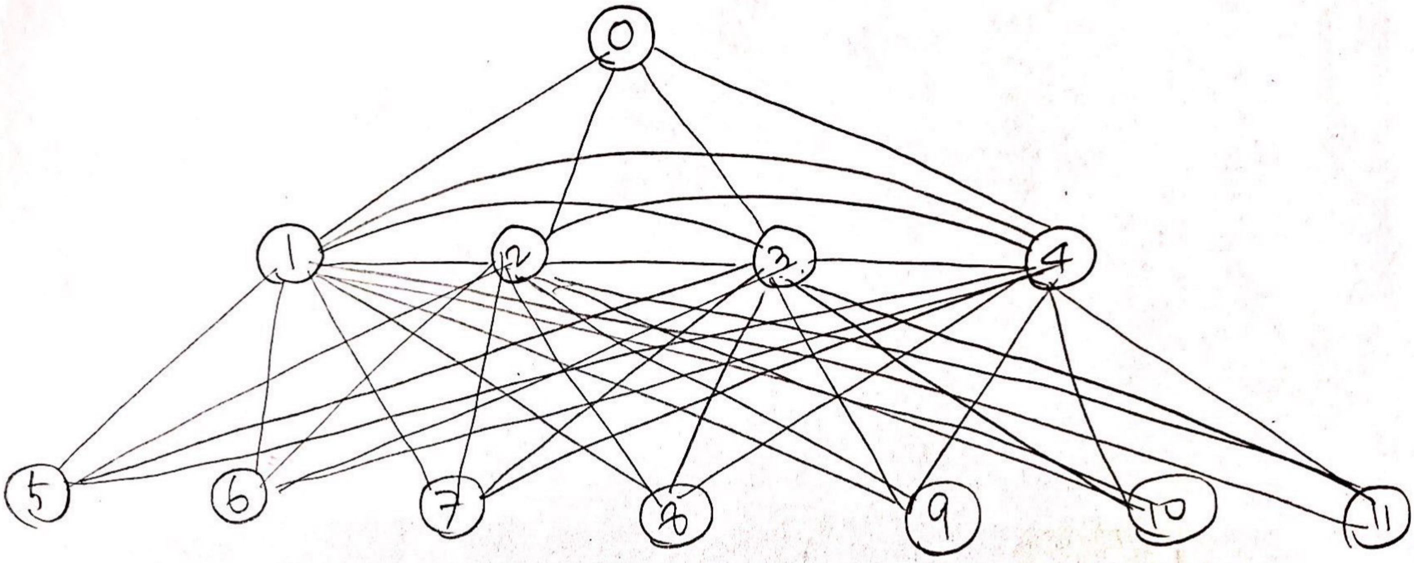
If we $\boxed{\text{sum over } y}$, the remaining variables $x_1, \cdots, x_T$ make a linear chain.

Then, $p(x_T)$ can be computed in linear time to T.

prob8.

a). Variable Elimination

The moralized graph is as follows.



A good elimination order is

11, 10, 9, 8, 7, 6, 5, 0, 4, 3, 2, 1.