Question 1

a. $P(C|r,w,s)$
$= \alpha P(C,r,w,s)$
$= \alpha P(C)P(r|C)P(s|C)P(w|s,r)$
$= \alpha\langle 0.5, 0.5\rangle\langle 0.8, 0.2\rangle\langle 0.1, 0.5\rangle \times 0.99$
$= \alpha\langle 0.0396, 0.0459\rangle$
$= \langle 0.4444, 0.5556\rangle$

$P(C|\neg r,w,s)$
$= \alpha P(C,\neg r,w,s)$
$= \alpha P(C)P(\neg r|C)P(s|C)P(w|s,\neg r)$
$= \alpha\langle 0.5, 0.5\rangle\langle 0.2, 0.8\rangle\langle 0.1, 0.5\rangle \times 0.90$
$= \alpha\langle 0.009, 0.18\rangle$
$= \langle 0.047619, 0.952381\rangle$

$P(R|c,w,s)$
$= \alpha P(c,R,w,s)$
$= \alpha P(c)P(R|c)P(s|c)P(w|s,R)$
$= \alpha 0.5 \times \langle 0.8, 0.2\rangle \times 0.1 \times \langle 0.99, 0.90\rangle$
$= \alpha\langle 0.0396, 0.009\rangle$
$= \langle 0.814815, 0.185185\rangle$

$P(R|\neg c,w,s)$
$= \alpha P(\neg c,R,w,s)$
$= \alpha P(\neg c)P(R|\neg c)P(s|\neg c)P(w|s,R)$
$= \alpha 0.5 \times \langle 0.2, 0.8\rangle \times 0.5 \times \langle 0.99, 0.90\rangle$
$= \alpha\langle 0.0495, 0.18\rangle$
$= \langle 0.215686, 0.784314\rangle$
PS: The first part is assigned as **True** and the second **False**

b. Python code has be independently created.
One possible output is : $P(r|s,w) = 0.31926$

Question 2

a. See fig 1

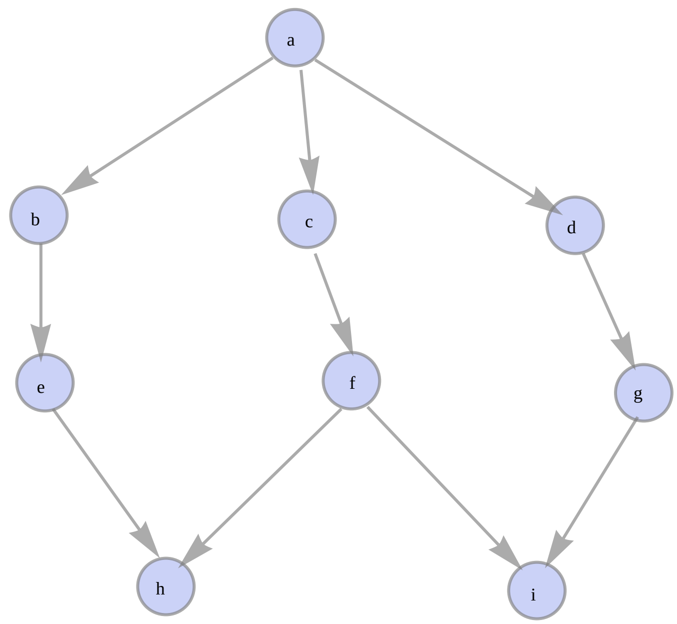

FIGURE 1. Belief Network

b. See fig 2

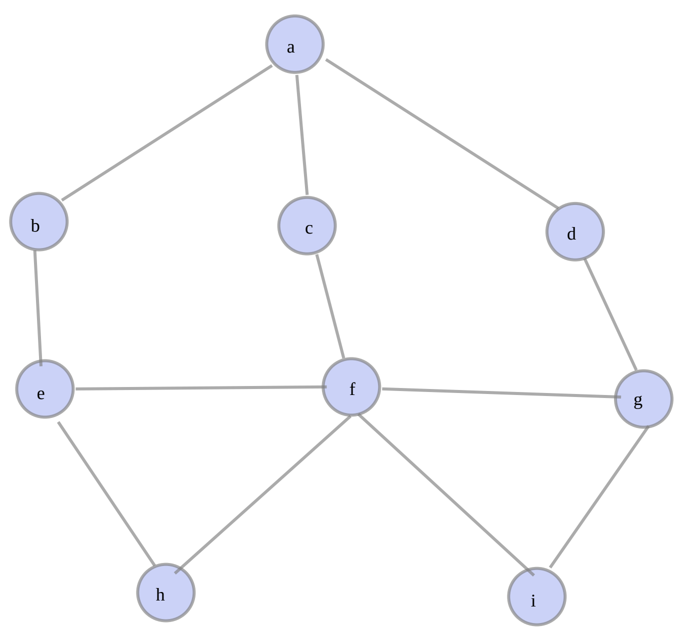

FIGURE 2. Moralized Graph

c. See fig 3

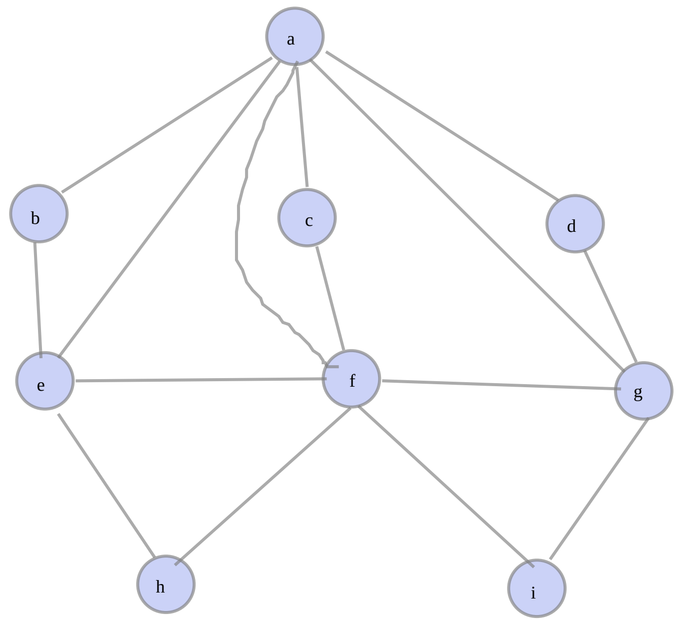

FIGURE 3. Triangulated Graph

d. See fig 4

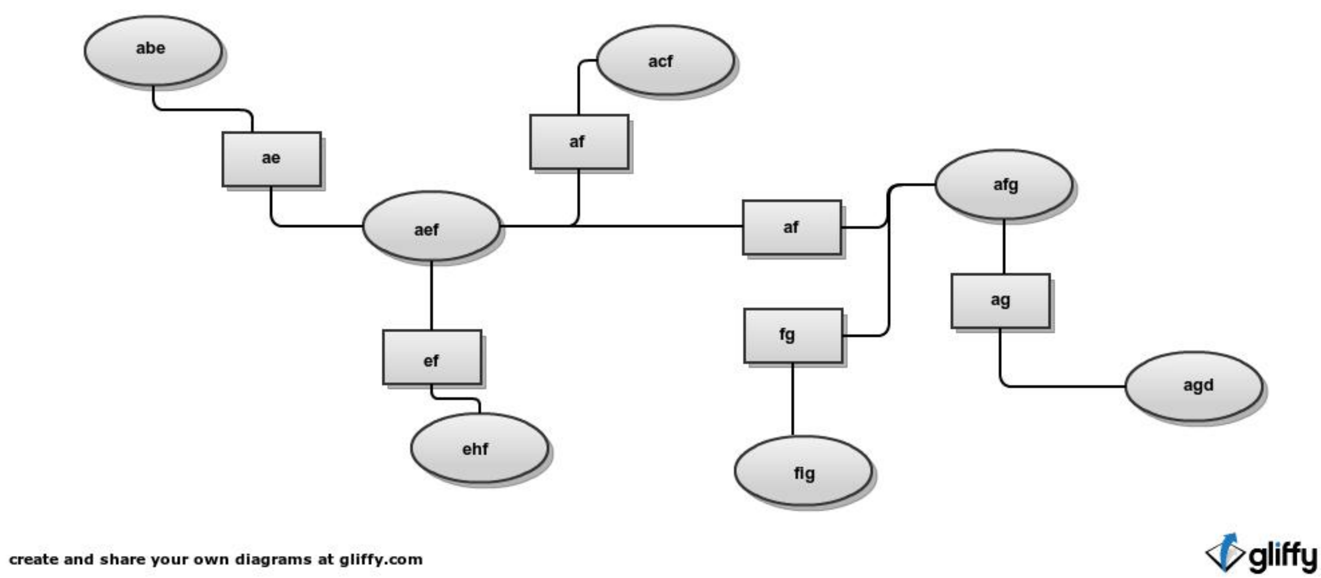

FIGURE 4. Junction Tree

e. $\Phi(abe) = p(a)p(b|a)p(e|b)$
$\Phi(aef) = 1$
$\Phi(acf) = p(c|a)p(f|c)$
$\Phi(afg) = 1$
$\Phi(agd) = p(d|a)p(g|d)$
$\Phi(ehf) = p(h|e, f)$
$\Phi(fig) = p(i|f, g)$
and all potentials of separators are set to be 1.

f.

(1) $abe \to aef: \ \Phi^*(ae) = \sum_b \Phi(abe) = p(a) \sum_b p(b|a)p(e|b)$
(2) $ehf \to aef: \ \Phi^*(ef) = \sum_h \Phi(ehf) = \sum_h p(h|e, f) = 1$
(3) $acf \to aef: \ \Phi^*(af) = \sum_c \Phi(acf) = \sum_c p(c|a)p(f|c)$
(4) $\Phi^*(aef) = \Phi(aef) \frac{\Phi^*(ae)\Phi^*(ef)\Phi^*(af)}{\Phi(ae)\Phi(ef)\Phi(af)}$
(5) $agd \to afg: \ \Phi^*(ag) = \sum_d \Phi(agd) = \sum_d p(d|a)p(g|d)$
(6) $fig \to afg: \ \Phi^*(fg) = \sum_i \Phi(fig) = \sum_i p(i|f, g) = 1$
(7) $\Phi^*(afg) = \Phi(afg) \frac{\Phi^*(ag)\Phi^*(fg)}{\Phi(ag)\Phi(fg)}$
(8) $afg \to aef: \ \Phi^{**}(af) = \sum_e \Phi^*(afg)$
(9) $\Phi^{**}(aef) = \Phi^*(aef) \frac{\Phi^{**}(af)}{\Phi^*(af)}$
(10) $abe \leftarrow aef: \ \Phi^{**}(ae) = \sum_f \Phi^{**}(aef)$
(11) $ehf \leftarrow aef: \ \Phi^{**}(ef) = \sum_a \Phi^{**}(aef)$
(12) $acf \leftarrow aef: \ \Phi^{**}(af) = \sum_e \Phi^{**}(aef)$
(13) $\Phi^{**}(abe) = \Phi(abe) \frac{\Phi^{**}(ae)}{\Phi^*(ae)}$
(14) $\Phi^{**}(ehf) = \Phi(ehf) \frac{\Phi^{**}(ef)}{\Phi^*(ef)}$
(15) $\Phi^{**}(acf) = \Phi(acf) \frac{\Phi^{**}(af)}{\Phi^*(af)}$
(16) $aef \to afg: \ \Phi^{**}(afg) = \dots$
(17) $afg \to agd, afg \to agd$ I'm tired...

Question 3

a. Since $x_{T+1}, x_{T+2}, \dots x_T$ do not matter in this question, I have omitted them. Bayesian Networks as below:

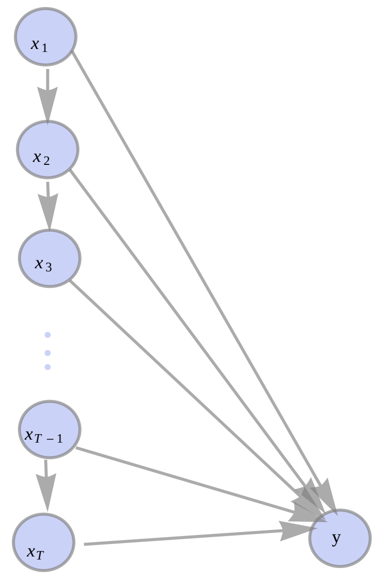

FIGURE 5. Bayesian Networks

Well, the corresponding Junction Tree is pretty complex, I'm not gonna to show a complete figure of it, instead, I would describe it and give a simple illustration when $T = 3$.

Description: *Through the step of* ***Moralisation****, since* $x_1, x_2, \ldots x_T$ *are parents of* y*, all of them should be connected each other. After Moralisation, there is no need to do Triangulation, there are* $\begin{pmatrix} T+1 \\ 3 \end{pmatrix}$ *cliques here, Thus we have* $\begin{pmatrix} T+1 \\ 3 \end{pmatrix} - 1$ *separators in the* ***Junction Tree****, consider the process of* ***Message Passing****, each separator has to be passed twice(two direction), we know that the computation of* $P(x_T)$ *requires* $O(T^3)$ *in this case.*

Following picture is a Junction Tree when $T = 3$, $\begin{pmatrix} 3+1 \\ 3 \end{pmatrix} = 4$ thus we have 4 cliques.

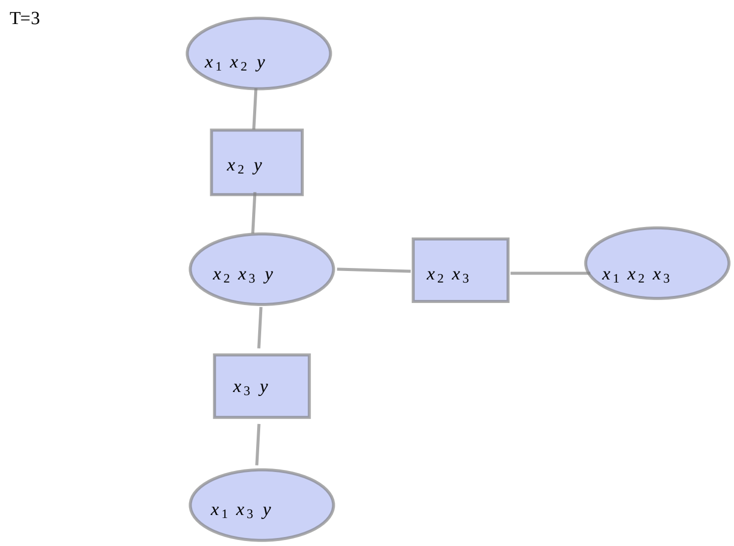

Figure 6. Junction Tree

b. Rewrite the following formula

$$p(x_T) = \sum_{y, x_1, x_2, \ldots, x_{T-1}} p(y|x_1, x_2, \ldots, x_T) p(x_1) \prod_{t=2}^{T} p(x_t|x_{t-1})$$

we get

$$p(x_T) = \sum_{y, x_1, x_2, \ldots, x_{T-1}} p(x_1) \prod_{t=2}^{T} p(x_t|x_{t-1}) \sum_{y} p(y|x_1, x_2, \ldots, x_T)$$

Note that

$$\sum_{y} p(y|x_1, x_2, \ldots, x_T) = 1$$

we get

$$p(x_T) = \sum_{y, x_1, x_2, \ldots, x_{T-1}} p(x_1) \prod_{t=2}^{T} p(x_t|x_{t-1})$$

Which implies

$$p(x_T) = \sum_{x_{T-1}} p(x_T|x_{T-1}) p(x_{T-1})$$

We first calculate the table of $p(x_2)$, which requires $O(1)$, then using $p(x_2)$ to calculate the table of $p(x_3)$, which also requires $O(1)$,..., finally we get the result of $p(x_T)$, it's trivial to see that the overall time complexity is $O(T)$.

Question 4

a. run **SmoothHMM.py** for both evidence sequences, we got:

(1) $e_{1:10} = (F, F, F, T, T, T, T, F, F, F)$

```
P(X_1=T|e1:10) = 0.01594
P(X_2=T|e1:10) = 0.02273
P(X_3=T|e1:10) = 0.31481
P(X_4=T|e1:10) = 0.90293
P(X_5=T|e1:10) = 0.94193
P(X_6=T|e1:10) = 0.90294
P(X_7=T|e1:10) = 0.31514
P(X_8=T|e1:10) = 0.02304
P(X_9=T|e1:10) = 0.01736
P(X_10=T|e1:10) = 0.09837
```

(2) $e_{1:10} = (F, T, F, T, F, T, F, T, F, T)$

```
P(X_1=T|e1:10) = 0.24747
P(X_2=T|e1:10) = 0.29091
P(X_3=T|e1:10) = 0.29778
P(X_4=T|e1:10) = 0.29982
P(X_5=T|e1:10) = 0.30030
P(X_6=T|e1:10) = 0.30130
P(X_7=T|e1:10) = 0.30454
P(X_8=T|e1:10) = 0.32356
P(X_9=T|e1:10) = 0.40594
P(X_10=T|e1:10) = 0.71520
```

b. run **MaxSeq.py**, we got:

(1) $e_{1:10} = (F, F, F, T, T, T, T, F, F, F)$

```
[0, 0, 0, 1, 1, 1, 1, 0, 0, 0]
```

(2) $e_{1:10} = (F, T, F, T, F, T, F, T, F, T)$

```
[0, 1, 0, 1, 0, 1, 0, 1, 0, 1]
```