

ex6.7:

The complexity of computing the normalisation constant:

$$T = O(2^{2n-1} \times (n - 1)) = O(n \cdot 2^{2n})$$

The normalisation constant $\log_n = 188.4774$ (The code for Q6.7 is in Appendix 'computeNormalisationConstant.m')

ex6.9

1.

The marginals of the symptoms is showed below, and the code for this question is in Appendix 'q1_JT_diseaseNet.m').

$$P(s1=1) = 0.44183$$

$$P(s2=1) = 0.45668$$

$$P(s3=1) = 0.44141$$

$$P(s4=1) = 0.49127$$

$$P(s5=1) = 0.49389$$

$$P(s6=1) = 0.65748$$

$$P(s7=1) = 0.50456$$

$$P(s8=1) = 0.26869$$

$$P(s9=1) = 0.64908$$

$$P(s10=1) = 0.49074$$

$$P(s11=1) = 0.42255$$

$$P(s12=1) = 0.4291$$

$$P(s13=1) = 0.54502$$

$$P(s14=1) = 0.63296$$

$$P(s15=1) = 0.42954$$

$$P(s16=1) = 0.45879$$

$$P(s17=1) = 0.42756$$

$$P(s18=1) = 0.40426$$

$$P(s19=1) = 0.58209$$

$$P(s20=1) = 0.58959$$

$$P(s21=1) = 0.76127$$

$$P(s22=1) = 0.69559$$

$$P(s23=1) = 0.5087$$

$$P(s24=1) = 0.41996$$

$$P(s25=1) = 0.35194$$

$$P(s26=1) = 0.38961$$

$$P(s27=1) = 0.32597$$

$$P(s28=1) = 0.46962$$

$$P(s29=1) = 0.52287$$

$$P(s30=1) = 0.71731$$

$$P(s31=1) = 0.5242$$

$$P(s32=1) = 0.3537$$

$$P(s33=1) = 0.51268$$

$$P(s34=1) = 0.5294$$

$$P(s35=1) = 0.38575$$

$P(s36=1) = 0.48909$
 $P(s37=1) = 0.6336$
 $P(s38=1) = 0.5896$
 $P(s39=1) = 0.42316$
 $P(s40=1) = 0.52823$

2.

Since each symptom connects to 3 parent diseases, which is fixed, we can calculate $P(s_i=1)$ by the basic Bayes' Rule:

$$p(s_i = 1, pa(s_i)) = p(s_i|pa(s_i)) \cdot p(pa(s_i)) \ \& \ p(s_i = 1) = \sum_{pa(s_i)} p(s_i = 1, pa(s_i))$$

and the result showed below is the same with that using the Junction Tree, and the code for this is 'q2_efficient_diseaseNet.m':

$P(s1=1) = 0.44183$
 $P(s2=1) = 0.45668$
 $P(s3=1) = 0.44141$
 $P(s4=1) = 0.49127$
 $P(s5=1) = 0.49389$
 $P(s6=1) = 0.65748$
 $P(s7=1) = 0.50456$
 $P(s8=1) = 0.26869$
 $P(s9=1) = 0.64908$
 $P(s10=1) = 0.49074$
 $P(s11=1) = 0.42255$
 $P(s12=1) = 0.4291$
 $P(s13=1) = 0.54502$
 $P(s14=1) = 0.63296$
 $P(s15=1) = 0.42954$
 $P(s16=1) = 0.45879$
 $P(s17=1) = 0.42756$
 $P(s18=1) = 0.40426$
 $P(s19=1) = 0.58209$
 $P(s20=1) = 0.58959$
 $P(s21=1) = 0.76127$
 $P(s22=1) = 0.69559$
 $P(s23=1) = 0.5087$
 $P(s24=1) = 0.41996$
 $P(s25=1) = 0.35194$
 $P(s26=1) = 0.38961$
 $P(s27=1) = 0.32597$
 $P(s28=1) = 0.46962$
 $P(s29=1) = 0.52287$
 $P(s30=1) = 0.71731$
 $P(s31=1) = 0.5242$
 $P(s32=1) = 0.3537$

$P(s_{33}=1) = 0.51268$
 $P(s_{34}=1) = 0.5294$
 $P(s_{35}=1) = 0.38575$
 $P(s_{36}=1) = 0.48909$
 $P(s_{37}=1) = 0.6336$
 $P(s_{38}=1) = 0.5896$
 $P(s_{39}=1) = 0.42316$
 $P(s_{40}=1) = 0.52823$

By this method, the complexity to calculate marginals is $T = O(2^3)$. Using Junction Tree, since each s_i has three parents d , a clique contains at least 4 variables, leading to the minimal computational complexity $O(2^3)$, but it can be larger. Therefore, the latter method is more efficient.

3.

the marginal for all diseases is showed below, and its code is 'q3_computeDisease.m', which is in the Appendix.

$p(d_1=1|s_{1:10}) = 0.029776$
 $p(d_2=1|s_{1:10}) = 0.38176$
 $p(d_3=1|s_{1:10}) = 0.95423$
 $p(d_4=1|s_{1:10}) = 0.39664$
 $p(d_5=1|s_{1:10}) = 0.49647$
 $p(d_6=1|s_{1:10}) = 0.43515$
 $p(d_7=1|s_{1:10}) = 0.18749$
 $p(d_8=1|s_{1:10}) = 0.70118$
 $p(d_9=1|s_{1:10}) = 0.043127$
 $p(d_{10}=1|s_{1:10}) = 0.61031$
 $p(d_{11}=1|s_{1:10}) = 0.28732$
 $p(d_{12}=1|s_{1:10}) = 0.48983$
 $p(d_{13}=1|s_{1:10}) = 0.8996$
 $p(d_{14}=1|s_{1:10}) = 0.61956$
 $p(d_{15}=1|s_{1:10}) = 0.92048$
 $p(d_{16}=1|s_{1:10}) = 0.7061$
 $p(d_{17}=1|s_{1:10}) = 0.20125$
 $p(d_{18}=1|s_{1:10}) = 0.90849$
 $p(d_{19}=1|s_{1:10}) = 0.86497$
 $p(d_{20}=1|s_{1:10}) = 0.88393$

ex6.10

1.

[[[[[[[ex6.10_1.jpg]]]]]]]]

Since all the variables are binary, the computational complexity of computing $p(x_T)$ is $O(2^T)$

2.

$$p(x_T) = \sum_{x_1} p(x_1) \cdot \sum_{x_2} p(x_2|x_1) \cdot \dots \cdot \sum_{x_T} p(x_T|x_{T-1}) \cdot \sum_y p(y|x_1, x_2, \dots, x_T)$$

By passing messages in this way, the complexity of computing $p(x_T)$ is $O(2 \times T) = O(T)$