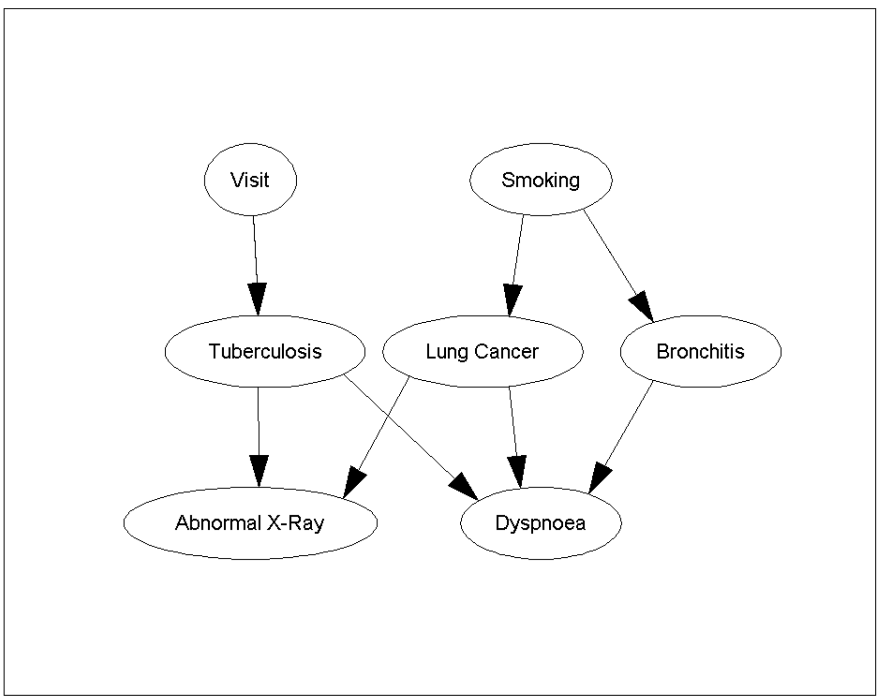
CPSC 540 Assignment 8 (due December 17)

Graphical Models

Please put your name and student number on the assignment and staple the assignment together. There is a +1 bonus if the assignment is typed in LaTeX. Please submit the assignment to me directly in ICCS 193, or if I am not around you can place it in my mailbox in ICCS 201 during normal business hours. Keep in mind that only the top 6 assignments count, so if you are happy with your mark on six of the first seven assignments there is no need to do this assignment. Because this is the end of term, early assignments are welcome but late assignments will not be accepted.

1 Conditional Independence

Consider the Bayesian network below, for distinguishing between different causes of shortness-of-breath (dyspnoea) and the causes of an abnormal lung x-ray, while modelling potential causes of these diseases too (whether the preson is a smoker or had a 'visit' to a country with a high degree to tuberculosis).



Assuming that all conditional independence properties of the distribution are reflected in the graph structure, say why each of the following conditional independence statements is true or false:

1. (Smoking) $\perp$ (Lung Cancer).
2. (Smoking) $\perp$ (Abnormal X-Ray).
3. (Lung Cancer) $\perp$ (Bronchitis).
4. (Tuberculosis $\perp$ (Bronchitis).
5. (Visit) $\perp$ (Abnormal X-Ray) | (Tuberculosis).
6. (Lung Cancer) $\perp$ (Bronchitis) | (Smoking).
7. (Tuberculosis) $\perp$ (Bronchitis) | (Dyspnoea).
8. (Tuberculosis) $\perp$ (Bronchitis) | (Abnormal X-Ray).
9. (Visit) $\perp$ (Smoking) | (Dyspnoea).
10. (Visit,Smoking) $\perp$ (Abnormal X-Ray, Dyspnoea) | (Tuberculosis, Lung Cancer, Bronchitis)

2 Exact Inference

Assume that we use the following parameterization of the network:

$$\begin{aligned}
&\text{Visit}\\
p(V=1) &= 0.01\\
&\text{Smoking}\\
p(S=1) &= 0.2\\
&\text{Tuberculosis}\\
p(T=1|V=1) &= 0.05\\
p(T=1|V=0) &= 0.01\\
&\text{Lung Cancer}\\
p(L=1|S=1) &= 0.10\\
p(L=1|S=0) &= 0.01\\
&\text{Bronchitis}\\
p(B=1|S=1) &= 0.60\\
p(B=1|S=0) &= 0.30\\
&\text{Abnormal X-Ray}\\
p(X=1|T=1,L=1) &= 1.00\\
p(X=1|T=1,L=0) &= 0.98\\
p(X=1|T=0,L=1) &= 0.9\\
p(X=1|T=0,L=0) &= 0.05
\end{aligned}$$

Dyspnoea

$$p(D = 1|T = 1, L = 1, B = 1) = 0.90$$
$$p(D = 1|T = 1, L = 1, B = 0) = 0.70$$
$$p(D = 1|T = 1, L = 0, B = 1) = 0.85$$
$$p(D = 1|T = 1, L = 0, B = 0) = 0.65$$
$$p(D = 1|T = 0, L = 1, B = 1) = 0.82$$
$$p(D = 1|T = 0, L = 1, B = 0) = 0.60$$
$$p(D = 1|T = 0, L = 0, B = 1) = 0.80$$
$$p(D = 1|T = 0, L = 0, B = 0) = 0.10$$

Compute the following quantities (hints are given on the right, and these will be easier to do in order and if you use conditional independence properties to simplify the calculations):

0. $p(S = 1)$ (marginal of root node; read from table)
1. $p(S = 0)$ (negation of marginal of root node; use sum to one constraint)
2. $p(L = 1|S = 1)$ (conditional of child node given parents; read from table)
3. $p(L = 1)$ (marginal of child node; marginalize over parent)
4. $p(X = 1|T = 1, L = 1)$ (conditional of child given parents; read from table)
5. $p(X = 1|T = 1)$ (conditional of child with missing parent; marginalize over missing parent)
6. $p(X = 1|T = 1, S = 1)$ (conditional of child given parent and grand-parent, marginalize over missing parent)
7. $p(X = 1)$ (marginal of leaf node; marginalize over parents and use independence to simplify)
8. $p(T = 1|X = 1)$ (conditional of parent given child; use Bayes rule)
9. $p(T = 1|L = 1)$ (conditional of parent given co-parent; use independence and then marginal)
10. $p(T = 1|X = 1, L = 1)$ (conditional of parent given child and co-parent; use Bayes rule)

3 Viterbi Decoding

The function *viterbiDemo.m* loads the initial state probabilities and transition probabilities for three Markov chain models on T binary variables,

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

It then tries to find the optimal decoding (the most likely assignment to the variables $\{x_1, x_2, \ldots, x_T\}$) in each of the three chains. In the demo, decoding is done by enumerating all possible assignments to the variables. This works for the first two chains as they only have 4 variables, but is too slow on the last chain because it has 30 variables.

Write a function, *viterbiDecode.m*, that implements the Viterbi decoding algorithm for Markov chains. You should use a 2 by T matrix V to represent the dynamic programming table, and another matrix T containing the argmax that lead to each entry in the table. Hand in this code and report the optimal decoding of the third Markov chain.