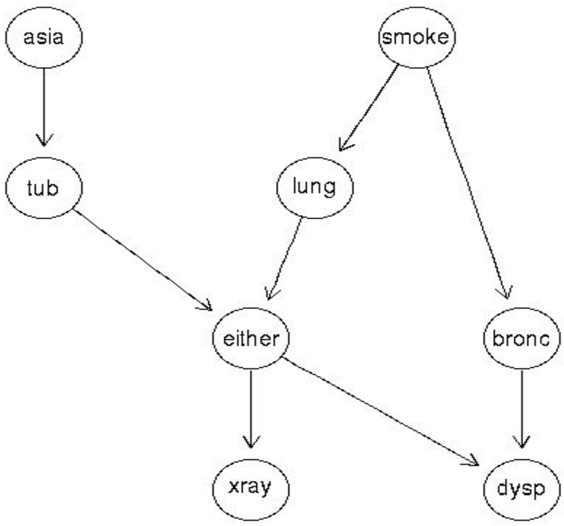
Problem Set 3: Exact Inference with Probabilistic Graphical Models

-By Rhea Carmel Glen Rodrigues

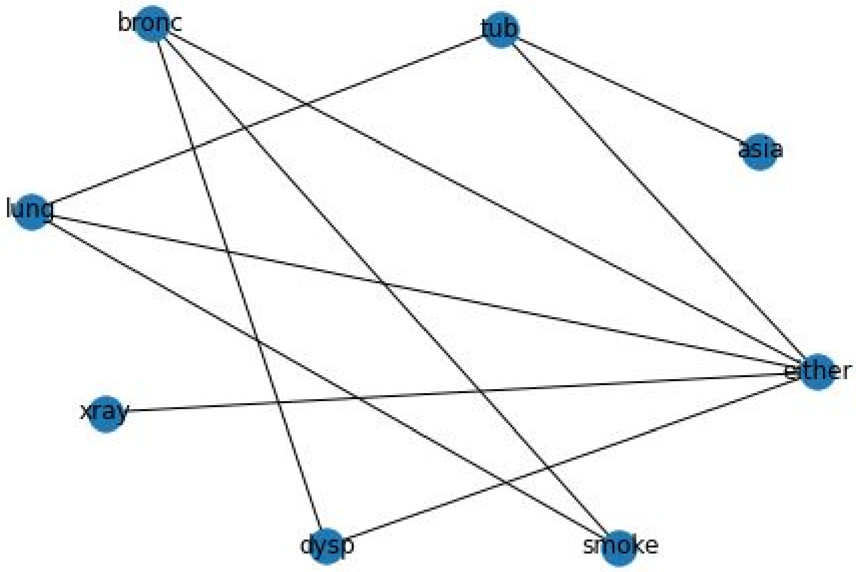
Make exact inferences about probabilistic graphical models using the state-of-the-art graphical model packages in our most comfortable programming languages, and understand those exact algorithms.

Work with the chest clinic graphical model (below). Please moralize, triangulate and construct a junction tree from this graphical model. Then use a message-passing algorithm to find the joint probability of "tub=yes, lung=yes, bronc=yes", given evidence that "asia=yes, xray=yes".



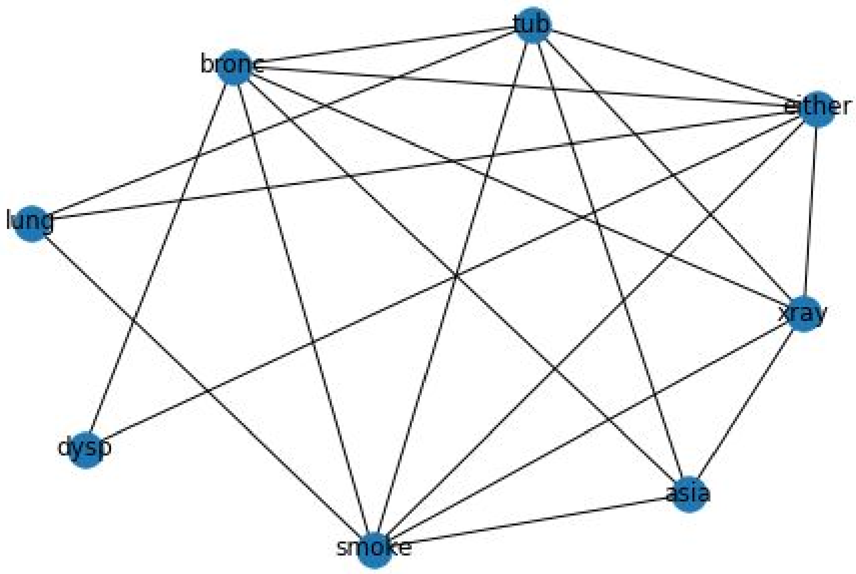
Draw the moral graph, triangulated graph and the junction tree.

The moralization converts a bayesian network into an undirected graphical model. That is if any pair of nodes have a common child then the nodes are connected . In our example the moralized graph looks as follows.



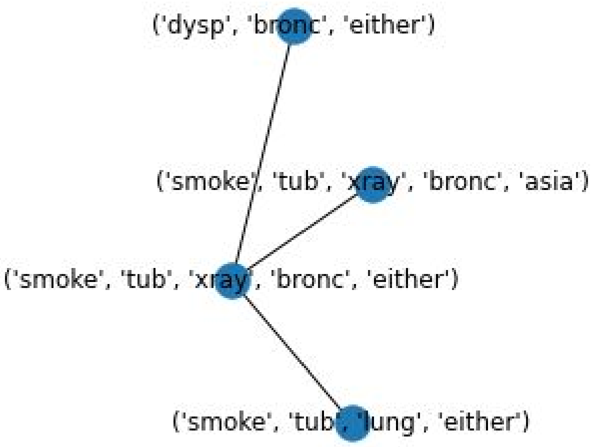
Triangulated graph:

In this graph there are no cycles of nodes whose length is larger than 4, which means any cycles with length larger than 3 needs to add a chord.



Junction tree.

Junction trees are created using clique trees which only have a maximal spanning tree. The junction tree is formed from triangulated graphs. The junction tree produced by using pgmpy .to_junction_tree() is as below.



The joint probability of "tub=yes, lung=yes, bronc=yes", given evidence that "asia=yes, xray=yes" is 0.016

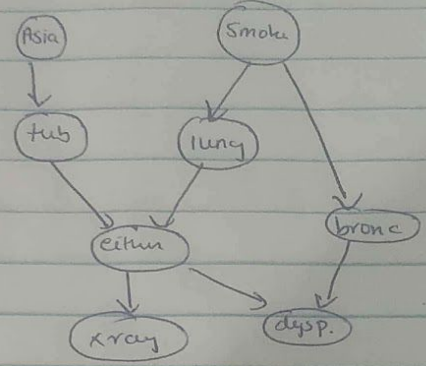
Explain why the "running intersection property" is satisfied in your junction tree.

According to the running intersection property, if a variable X is present in two clusters it should also be present in the unique path that connects the two clusters. In our example cluster containing ('dspy','bronc','either') and ('smoke','tub','lung','either'') each have variable 'either'' . The two clusters are connected with the cluster('smoke','tub','xray','bronc','either') which also has a variable 'either'. hence it follows the property of running the intersection property.

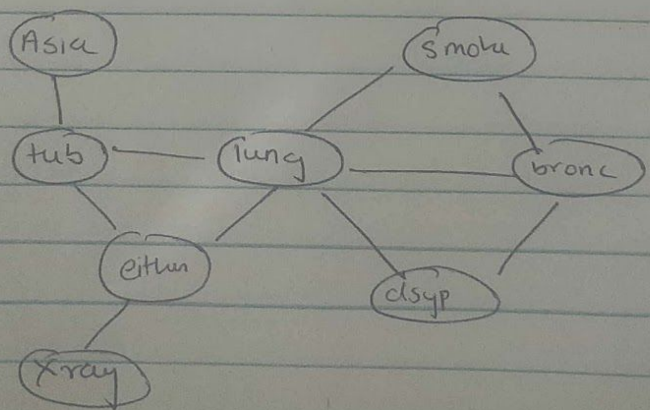
Problem 2

Describe how the different terms on the right hand side of "p(V) = p(a)p(t | a)p(s)p(l | s)p(b | s)p(e | t, l)p(d | e, b)p(x | e)" are distributed among the different junction tree clusters. Write out the messages using these terms and verify that the message passing algorithm indeed gives the cluster marginals.

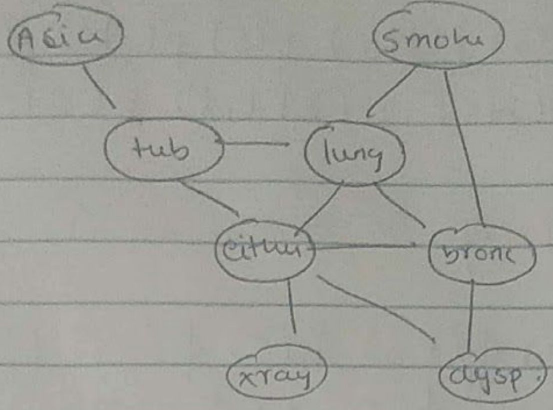
Q2 The graph provided is



The Moralization converts a bayesian network into undirected graphical model.
That is if any pair of Node have common child then the nodes are connected in our example. the ~~moralization~~ moralized graph looks like.



The Triangulated graph is



Junction Tree :-

Junction tree can be created by using variable elimination algorithm.
The Joint distribution is given as

$$P(v) = p(a)\,p(t|a)\,p(s)\,p(l|s)\,p(b|s)\,p(e|t,l)\,p(d|e,b)\,p(x|e)$$

where $a \to$ asia, $t \to$ tub, $s \to$ smoke, $l \to$ lung, $b \to$ bronc, $d \to$ dysp, $x \to$ xray, $e \to$ either

$P(v)$ can also be written with respect to factor Φ
where Φ is $\{ P(X_i | Pa_{X_i}) \}_{i=1}^{n}$ where X is the set of variable.
for example $\Phi_T(t,a) = P(a)\,p(t|a)$ —— (1)

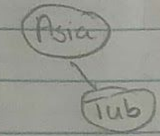
During variable elimination the order of the variable plays an important role.
We will be ~~using~~ following an order

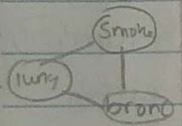
$$\sum_{e} \dots \sum_{a} p(a)\, p(t|a)\, p(s)\, p(l|s)\, \cancel{p(e|t,l)}$$
$$p(b|s)\ p(e|t,l)\ p(d|e,b)\ p(x|e)$$

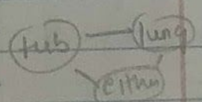
by the property as shown in 1 we get

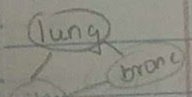
$$P(v) = \phi_A(a)\ \phi_T(t,a)\ \phi_S(s)\ \phi_L(l,s)\ \phi_E(e,t,l)$$
$$\phi_B(b,s)\ \phi_D(d,e,b)\ \phi_X(x,e)$$

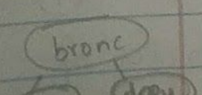
Using variable elimination:-







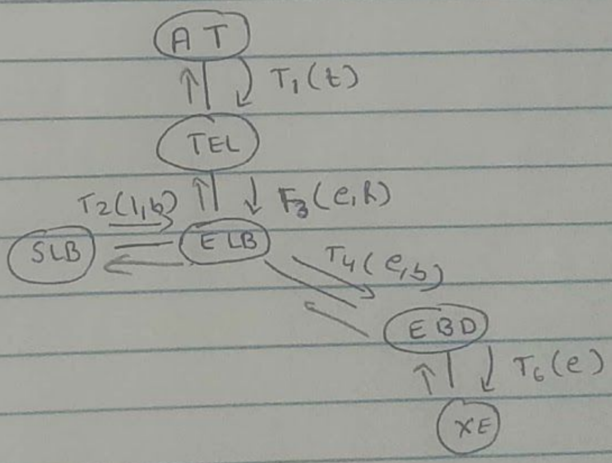


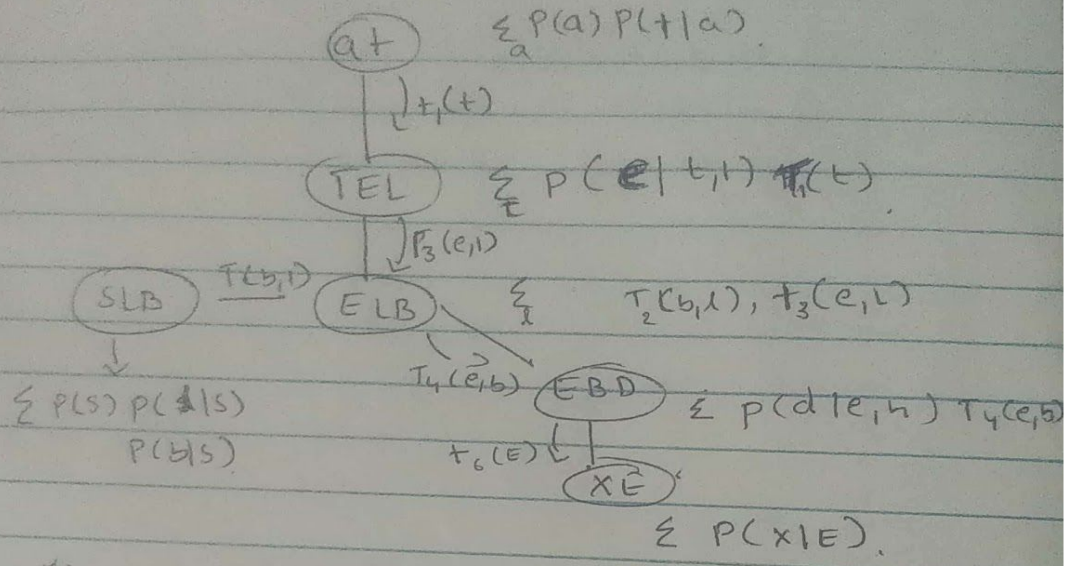


Variable	Factor Used.	variable involved	New factor
Asia.	$\phi_A(a)\ \phi_T(t,a)$	T, A.	$T_1(t)$
Smoke. ~~Tub~~	$\phi(s)\ \phi(l,s)$ $\phi(b,s)$	S, L, B	$T_2(l,b)$
Tub	$\phi(e,t,l)$, $T_1(t)$	E T L	$T_3(e,l)$
lung.	$T_3(e,l)$ $T_2(l,b)$	E, L, B	$T_4(e,b)$
Bronc	$T_4(e,b)$ $\phi(d,e,b)$	D E B	$T_5(DE)$

	Variable	Factor used	variable involved	New factor
either – Dspy	Dysp	$T_5(E,D)$	E D	$T_6(E)$
either – Xray	Either	$T_6(E)$ $\phi(x,e)$	E, X	$T_7(x)$

The Junction tree can be created using the variable involved as they act as ~~clique~~ a ~~variat~~ cluster.





The variable terms on right hand side are given as above.

Reference:

- https://github.com/pgmpy/pgmpy/blob/dev/examples/Creating%20a%20Bayesian%20Network.ipynb
- https://github.com/pgmpy/pgmpy/blob/dev/examples/Inference%20in%20Bayesian%20Networks.ipynb

- https://ublearns.blackboard.com/bbcswebdav/pid-5415559-dt-content-rid-30359140_1/courses/2201_19191_PsC/20200324-graphicalModels.html#(14)

- https://www.cs.toronto.edu/~urtasun/courses/GraphicalModels/lecture7.pdf

- https://learn-us-east-1-prod-fleet01-xythos.s3.us-east-1.amazonaws.com/5e00ea752296c/3876013?response-content-disposition=inline%3B%20filename%2A%3DUTF-8%27%27Building%2520Probabilistic%2520Graphical%2520Models%2520with%2520Python.pdf&response-content-type=application%2Fpdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Date=20200404T015809Z&X-Amz-SignedHeaders=host&X-Amz-Expires=21600&X-Amz-Credential=AKIAZH6WM4PLTYPZRQMY%2F20200404%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Signature=3e97bfaf17f0aaa9377158f1f3394f9db37f1eedf6376ec1534e44cd5697e9b7
- https://networkx.github.io/documentation/stable/reference/drawing.html