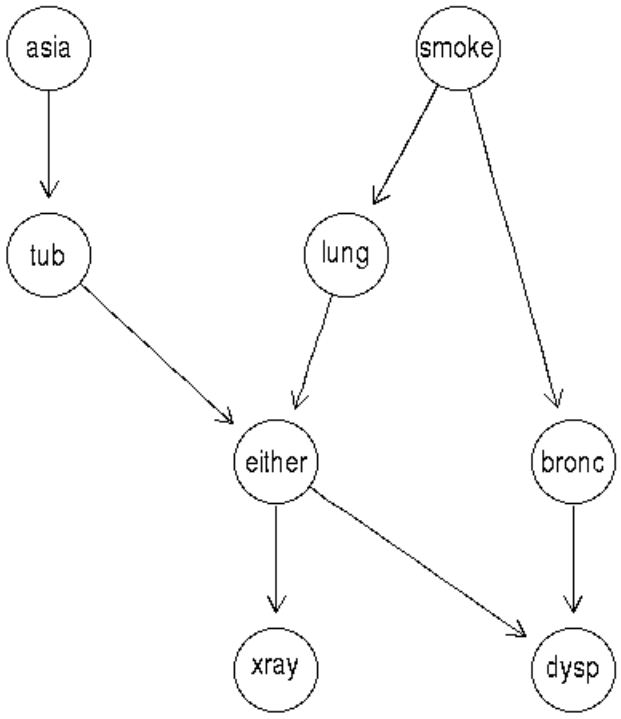
CSE455/555 - Intro to Pattern Recognition
Problem Set 3: Exact Inference with Probabilistic Graphical Models

Due Date: Friday, March 26, 2021 11:59PM

In this problem set, you will make exact inferences about probabilistic graphical models using the state-of-the-art graphical model packages in your most comfortable programming languages, and understand those exact algorithms. You can find tutorials in python, R (slides and book) and Matlab. Click the red words you will be directed to the related tutorials. If you can not open the first link, log into your UBLearn account and click the link again. The function calls in different packages are different, but the point here is that we make graphical model our actionable machine learning tool in this course.
You will work with the chest clinic graphical model below. The data you are going to use is in the following tables. The R code used to generate the data is also provided.




```
> library(gRain)
> yn <- c("yes","no")
> a <- cptable(~asia, values=c(1,99), levels=yn)
> t.a <- cptable(~tub | asia, values=c(5,95,1,99), levels=yn)
```

```
>  s <- cptable(~smoke, values=c(5,5), levels=yn)
>  l.s <- cptable(~lung | smoke, values=c(1,9,1,99), levels=yn)
>  b.s <- cptable(~bronc | smoke, values=c(6,4,3,7), levels=yn)
>  e.lt <- cptable(~either | lung:tub,values=c(1,0,1,0,1,0,0,1),
   levels=yn)
>  x.e <- cptable(~xray | either, values=c(98,2,5,95), levels=yn)
>  d.be<-cptable(~dysp|bronc:either, values=c(9,1,7,3,8,2,1,9),
   levels=yn)
> cpt.list <- compileCPT(list(a, t.a, s, l.s, b.s, e.lt, x.e, d.be))
> cpt.list$asia
asia
 yes   no
0.01 0.99
> cpt.list$tub
       asia
tub     yes   no
  yes 0.05 0.01
  no  0.95 0.99
> cpt.list$smoke
smoke
yes  no
0.5 0.5
> cpt.list$lung
      smoke
lung   yes   no
  yes 0.1 0.01
  no  0.9 0.99
> cpt.list$bronc
      smoke
bronc yes  no
  yes 0.6 0.3
  no  0.4 0.7
> ftable(cpt.list$either,row.vars = 1)
       lung yes     no
       tub  yes no yes no
either
yes           1   1   1   0
no            0   0   0   1
> cpt.list$xray
       either
xray    yes   no
  yes 0.98 0.05
  no  0.02 0.95
```

asia	
yes	no
0.01	0.99

	asia	
tub	yes	no
yes	0.05	0.01
no	0.95	0.99

smoke	
yes	no
0.5	0.5

	smoke	
lung	yes	no
yes	0.1	0.01
no	0.9	0.99

	smoke	
bronc	yes	no
yes	0.6	0.3
no	0.4	0.7

	lung	yes		no	
	tub	yes	no	yes	no
either					
yes		1	1	1	0
no		0	0	0	1

	either	
xray	yes	no
yes	0.98	0.05
no	0.02	0.95

1 Task-1

(a) Draw the moral graph, triangulated graph and the junction tree.
Please read the related parts in the tutorials and draw graphs using your choice of language. This is a coding task and your code will be graded. DO NOT draw graphs by hand and upload images.
(b) Explain why the "running intersection property" is satisfied in your junction tree.

2 Task-2

(a) 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.
(b) Write out the messages using these terms and verify that the message passing algorithm indeed gives the cluster marginals.
(c) Use message-passing algorithm to find the joint probability of "tub=yes, lung=yes, bronc=yes", given evidence that "asia=yes, xray=yes".
This is also a coding task, use your choice of language and refer to the tutorials.

3 Task-3[Optional]

Find the joint probability with MCMC.

4 Submission

Submit your solutions as a single ipynb file through UBlearn. You can use Google Colab:
https://colab.research.google.com/notebooks/intro.ipynb
https://towardsdatascience.com/getting-started-with-google-colab-f2fff97f594c.
The ipynb file should include your code, execution results, any explanations and answers to the questions. Use text cells to answer questions and add explanations.
Markdown guide for text cells:

https://colab.research.google.com/notebooks/markdown_guide.ipynb#scrollTo=Lhfnlq1Surtk
https://colab.research.google.com/notebooks/basic_features_overview.ipynb#scrollTo=4hfV37gxpP_c
You can also add math to text cells using LaTeX. Just place the statement within a pair of $ signs. Please typeset your mathematics. Do not upload pictures of handwriting math formulas. Math typesetting help: https://www.codecogs.com/latex/eqneditor.php

5 Libraries

No limitation on libraries for this assignment.

6 Rubric

Total: 10 points + 2 bonus points

Task-1:
5 points: 4 points part(a), 1 points part(b).
Task-2:
5 points: 1 points part(a), 3 points part(b), 1 point part(c).
Task-3:
2 bonus points.

7 Late Penalty

Since this assignment is posted two days late, there will be no late penalty if you submit it within three days after the deadline (before March 29, 11:59 pm). 10% late penalty for submissions after three days but within a week after the deadline (before April 2, 11: 59 pm). Submissions more than a week late will not be graded. Please start early because you have to read tutorials for this assignment.

8 Acknowledgement

By submitting this paper, you agree: (1) that you are submitting your paper to be used and stored as part of the SafeAssign™ services in accordance with the Blackboard Privacy Policy; (2) that your institution may use your paper in accordance with your institution's policies; (3) that your use of SafeAssign will be without recourse against Blackboard Inc. and its affiliates.

9 Academic Integrity

Academic integrity is a fundamental university value. Any violation will be reported to the University and will result in penalties in grades.

Do not share your answers with other students. This is an individual assignment. You are not allowed to work in groups. Working in groups and submitting similar answers is considered a violation of academic integrity.
Do not plagiarise someone else's words, ideas, or data you find online. Always cite your sources.