CS 512: Advanced Machine Learning **Due: 12 noon, Feb 27, 2020**

Assignment 1: Graphical Models (Written Questions)

Student: Zhanibek Rysbek *Email: zrysbe2@uic.edu*

Deadline: 12 noon, Feb 27, 2020

This is the written part of Assignment 1, consisting of five questions. **They are for individual work; you must work on them by yourself.** You should write up your solution to the five questions in a single Portable Document Format (PDF) file, which can be either typed using WORD or latex, or be scanned copies of handwritten manuscripts provided that the handwriting is neat and legible. The LaTeX source code of this document is provided with the package, and you may write up your report based on it.

How to submit the written part solution. Name the PDF as `Surname_UIN.pdf`, where `Surname` is your last name and `UIN` is your UIC UIN. Upload the PDF to Blackboard, under `Assessment\Assignment_1_Written`.

Note there are two different links on Blackboard corresponding to the written and programming parts of Assignment 1.

You are allowed to resubmit as often as you like and your grade will be based on the last version submitted. Late submissions will not be accepted in any case, unless there is a documented personal emergency. Arrangements must be made with the instructor as soon as possible after the emergency arises, preferably well before the deadline. Assignment 1 contributes **13%** to your final grade, and the written part carries 45 points (out of 100) for Assignment 1.

Start working on the assignment early.

Latex primer: http://ctan.mackichan.com/info/lshort/english/lshort.pdf (Chapter 3)

Written Questions (45 pt, Work on Yourself Only)

1. (10 pt, Ex 2.1 of Murphy's book) My neighbor has two children. Assuming that the gender of a child is like a coin flip, it is most likely, a priori, that my neighbor has one boy and one girl, with probability 1/2. The other possibilities—two boys or two girls—have probabilities 1/4 and 1/4.

   (a) (2 pt) What is the a priori probability that exactly one child is a girl?

   (b) (2 pt) What is the a priori probability that at least one child is a girl?

   (c) (3 pt) Suppose I ask him whether he has any boys, and he says yes. What is the probability that (exactly) one child is a girl? How does it compare with the result of (a) and (b)?

   (d) (3 pt) Suppose instead that I happen to see one of his children run by, and it is a boy. What is the probability that the other child is a girl?

Solution:

(a) Sample space consists of three different possibility that neighbor might have two boys, one boy and one girl and two girls: $\Omega = \{BB, BG, GG\}$. Each of which has following priori probability:

$$P(BB) = \frac{1}{4}$$
$$P(GG) = \frac{1}{4}$$
$$P(BG) = \frac{1}{2}$$

Therefore, priori probability of having exactly one girl is $\frac{1}{2}$

(b) Similarly, priori probability of having at least one girl (denote it aG) is the sum of all probabilities having girl OR $1 - noG$:

$$P(aG) = P(GG) + P(BG) = \frac{3}{4}$$

(c) Assume that $eB, (eG)$ denotes the event of having exactly one boy (or girl), $aB, (aG)$ denotes of having at least one boy (girl). The question is to find the $P(eG|aB)$ - probability of exactly one girl given that neighbor has at least one boy.

$$P(eG|aB) = \frac{P(eG, aB)}{P(aB)} = \frac{P(BG)}{P(aB)} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3}$$

(d) The question is the probability of other child being girl. The statement of the question is asking about the gender of a child, which is assumed to be independent and like a flip of a coin. Therefore, answer is $\frac{1}{2}$:

$$P(1stG|otherB) = \frac{P(1stG, otherB)}{P(otherB)} = \frac{P((1stG)P(otherB))}{P(otherB)} = P(1stG) = \frac{1}{2}$$

2. (8 pt, Ex 2.2 of Murphy's book) Suppose a crime has been committed. Blood is found at the scene for which there is no innocent explanation. It is of a type which is present in 1% of the population.

(a) (3 pt) The prosecutor claims: There is a 1% chance that the defendant would have the crime blood type if he were innocent. Thus there is a 99% chance that he guilty. This is known as the prosecutor's fallacy. What is wrong with this argument?

(b) (5 pt) The defender claims: The crime occurred in a city of 800,000 people. The blood type would be found in approximately 8000 people. The evidence has provided a probability of just 1 in 8000 that the defendant is guilty, and thus has no relevance. This is known as the defender's fallacy. What is wrong with this argument?

Hint: write out the definition of events and formalize the probability in question.

Solution:

(a) Denote that event A is associated that defendant committed a crime. B - blood type found in scene matches with defendant's blood. Thus, $P(B) = 0.01$

What prosecutor is claiming is: $P(A = true|B = true) = 1 - P(B) = 0.99$. Apparently, it is not correct. What is supposed to be $P(A = true|B = true) = \frac{P(A=true,B=true)}{P(B=true)}$, which is hard to compute. Prosecutor is definitely exaggerating, and probability of 0.99 is too high.

(b) Defender is implicitly assuming that the event A and B has no relevance to each other which is too strong argument without evidence. In a mathematical terms: $A \perp B \implies P(A|B) = P(A)$. In reality, having two events happened usually increases the likelihood of joint probability (that defendant is guilty). More appropriate hypothesis would be $P(A|B) = \frac{P(A,B)}{P(B)}$.

3. (10 pt, Ex 2.9 of Murphy's book)
Are the following properties true? Prove or disprove. Note that we are not restricting attention to distributions that can be represented by a graphical model.

(a) (5 pt) True or false? $(X \perp W|Z, Y) \wedge (X \perp Y|Z) \Rightarrow (X \perp Y, W|Z)$.

(b) (5 pt) True or false? $(X \perp Y|Z) \wedge (X \perp Y|W) \Rightarrow (X \perp Y|Z, W)$.

Solution:

(a) The answer is true. Aforementioned statement could be proven using definition of conditional probability and chain rule.

Let's start with right hand side:

$$P(X,Y,W|Z) = \frac{P(X,Y,W,Z)}{P(Z)} = \frac{P(Z)P(X|Z)P(Y|X,Z)P(W|X,Y,Z)}{P(Z)} \tag{1}$$

$$= P(X|Z)P(Y|X,Z)P(W|X,Y,Z) \tag{2}$$

Using the conditional independence $X \perp Y|Z$ we simplify $P(Y|X,Z) = P(Y|Z)$. Similarly, $X \perp W|Z,Y$ implies $P(W|X,Y,Z) = P(W|Y,Z) = \frac{P(W,Y|Z)}{P(Y|Z)}$. Substituting these equations into 2 we get:

$$P(X,Y,W|Z) = P(X|Z)P(Y|Z)\frac{P(Y,W|Z)}{P(Y|Z)} = P(X|Z)P(W,Y|Z) \tag{3}$$

Therefore, it proves that $(X \perp Y,W|Z)$

(b) Answer is false. Consider following counter example. Let $X,Y,Z \in \{0,1\}$ be independent random variables each of them has equal likely of getting boolean values of 0 or 1. Finally, consider $W = X+Y+B$ be mutual exclusive OR of the rest of them. From given distribution, we can verify the conditional independence hold: $(X \perp Y|Z) \wedge (X \perp Y|W)$. Because, X and Y are independent anyways, conditioned on any other random variable still keeps the independence property.

However, $X \perp Y|Z,W$ is not true, because of the variable W, that forces X,Y to attain certain values that in combination will be equal to W. Therefore, this counter example disproves aforementioned statement.

4. (10 pt, Ex 2.6 of Murphy's book)
a. (5 pt) Let $H \in \{1,\dots,K\}$ be a discrete random variable, and let e_1 and e_2 be the observed values of two other random variables E_1 and E_2. Suppose we wish to calculate the vector

$$\overrightarrow{P}(H|e_1,e_2) = (P(H=1|e_1,e_2),\dots,P(H=K|e_1,e_2)).$$

Which of the following sets of numbers are sufficient for the calculation?

(i) $P(e_1,e_2), P(H), P(e_1|H), P(e_2|H)$.

(ii) $P(e_1,e_2), P(H), P(e_1,e_2|H)$.

(iii) $P(e_1|H), P(e_2|H), P(H)$.

b. (5 pt) Now suppose we now assume $E_1 \perp E_2|H$ (i.e., E_1 and E_2 are conditionally independent given H). Which of the above 3 sets are sufficient now?

Show your calculations as well as giving the final result. Hint: use Bayes rule.

Solution:

(a) Bayes rule: $P(H, e_1, e_2) = \frac{P(e_1,e_2|H)P(H)}{P(e_1,e_2)}$.

In fact, $P(e_1, e_2)$ could be computed using $P(H)$ and $P(e_1, e_2|H)$ only reducing the amount of information required:

$$P(e_1, e_2) = \sum_h P(e_1, e_2|H = h)P(H = h) \tag{4}$$

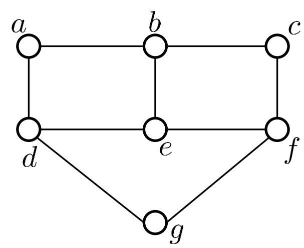
However, (ii) is the closest answer.

(b) Given that E_1 and E_2 is independent, one could compute the $\overrightarrow{P}$ using fewer amount of information:

$$P(H, e_1, e_2) = \frac{P(e_1, e_2|H)P(H)}{P(e_1, e_2)} = \frac{P(e_1|H)P(e_2|H)P(H)}{P(e_1, e_2)}$$

$P(e_1, e_2)$ is computed using equation 4. Therefore, closest answer is (iii).

5. (7 pt) What is the treewidth of the following graph? Write out all edges you add to triangulate the graph (*e.g.* (a,e)).

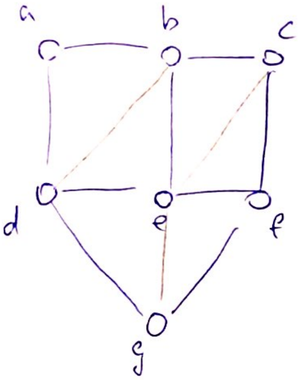


Solution:

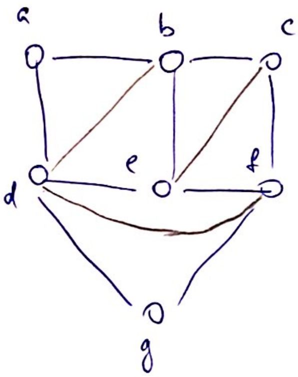
The graph contains 3 cycles with 4 edges $cycle_1 : (a, b, d, e), cycle_2 : (b, c, e, f), cycle_3 : (d.e, f, g)$. Each cycle has two possible option to triangulate $cycle_1 : (a, e), (b, d); cycle_2 : (b, f), (c, e); cycle_3 : (c, g), (d, f)$. So, amount of all possible triangulated graph is $2^3 = 8$.

Number of maximum clique for all possible graphs is 3 except the case when $(b, e), (b, f), (d, f)$ edges added making the max clique to be 4. Therefore, **treewidth** of the given graph is $3 - 1 = 2$.

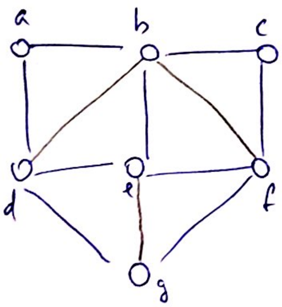
Refer to pictures below for more details.



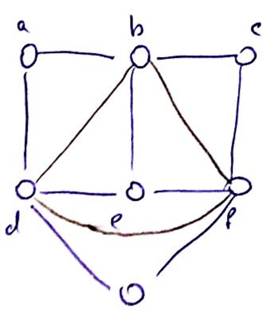
max clique: 3
{a, b, d} {b, d, e},
.. etc.



max clique: 3
(abd), (def), ..
etc.

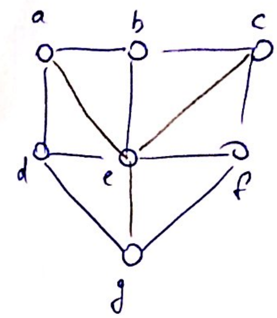


max clique: 3
abd, bde, bcf,
bef, ... etc.

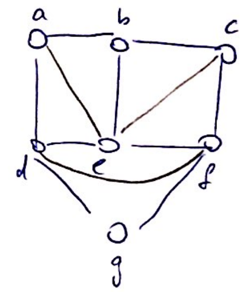


max clique: 4.
bdef, abd ...

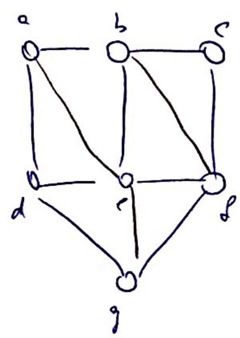
max clique: 3
abc, ...



max clique: 3
ade, ...



max clique: 3
deg, ..



max clique: 3
ade, etc.