ECE/CS 498 DSU/DSG Spring 2019
In-Class Activity 5

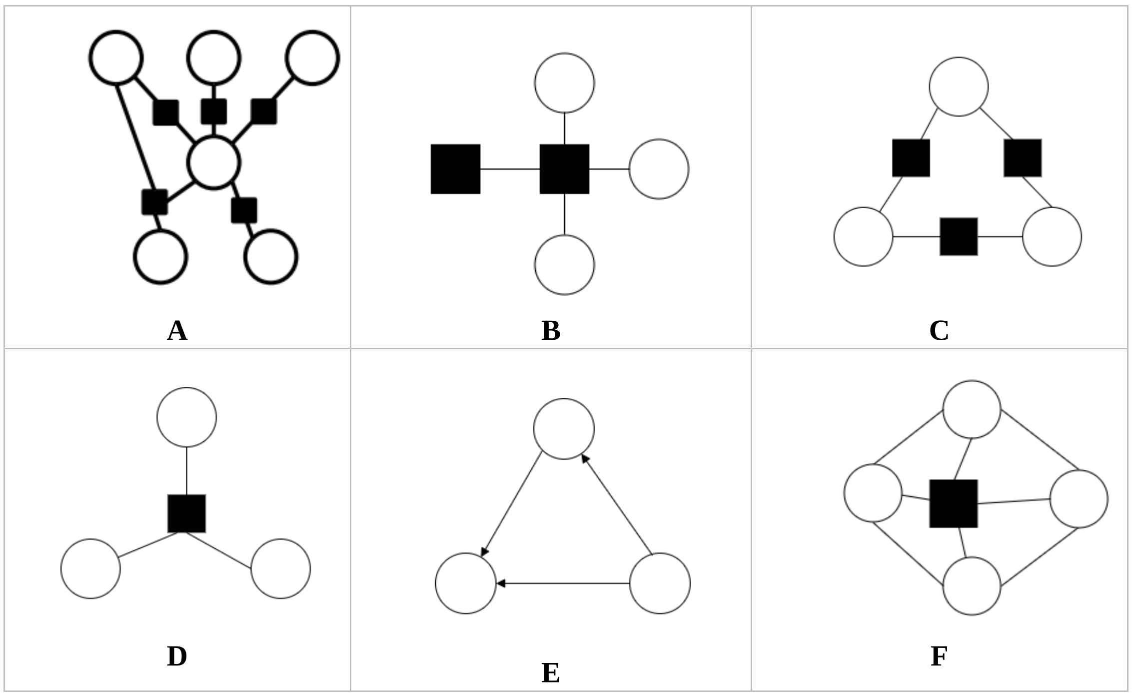
NetID: ____

The purpose of the in-class activity is for you to:

(i) Review concepts related to structure and conditional independence in factor graphs
(ii) Work out steps in belief propagation for inference on a factor graph

Factor Graphs

1. a) Which of the following graphical models are <u>invalid</u> representations of Factor Graphs? *(circle)* For invalid factor graphs, provide a short justification.



B, E and F are invalid factor graphs

B – Two factor functions connected to one another

E – Directed model. This is a valid Bayesian Network, but not a factor graph.

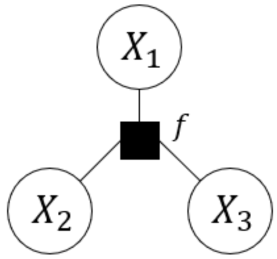
F – Variables connected without a factor function.

b) Consider random variables A, B connected by a factor function $f(A, B)$. Which of the following can $f(A, B)$ represent? (Mark all that apply)

i) Affinity between variables A and B (Yes)
ii) Conditional relation between variables A and B – for example $P(A|B)$ (Yes)
iii) Joint relation between variables A and B (Yes)

2. Given three binary random variables X_1, X_2, X_3 and a factor function $f(X_1, X_2, X_3)$

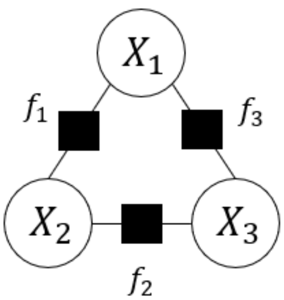
a) Draw and label the factor graph representing X_1, X_2, X_3 and f.



b) Assume that $f(X_1, X_2, X_3)$ factorizes to

$$f(X_1, X_2, X_3) = f_1(X_1, X_2) f_2(X_2, X_3) f_3(X_1, X_3)$$

Draw and label the factor graph.

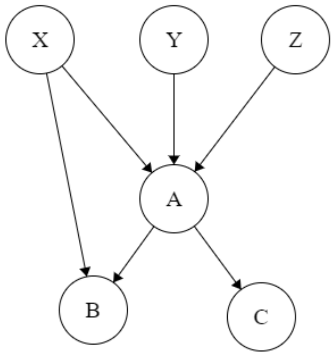


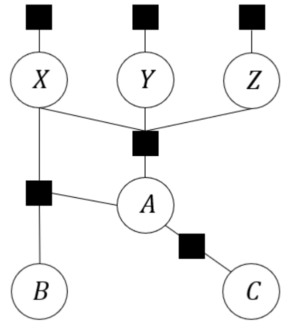
c) Compare the two factor graph models in a) and b) in terms of model complexity, i.e., number of parameters.

Model in a) requires specifying the function f which has 8 values (one for each combination of inputs). Therefore, it needs 8 parameters.

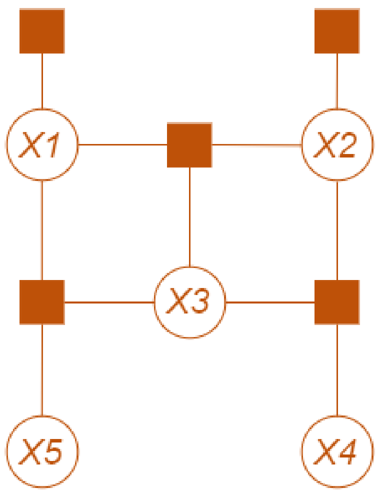
Model in b) requires specifying functions f_1, f_2, f_3, each of which has 4 values (one for each combination of inputs). Therefore, it needs 4x3=12 parameters.

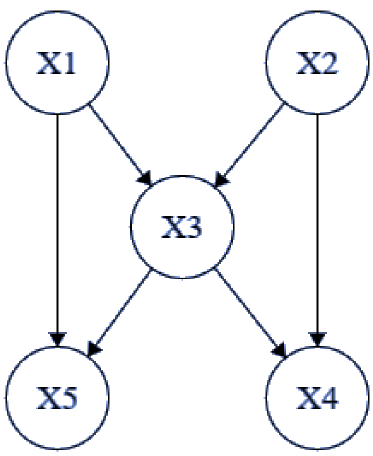
3. Bayesian Network and Factor Graph
   a) Convert the Bayesian Network below into a Factor Graph. Remember that each conditional distribution in the BN becomes a factor.



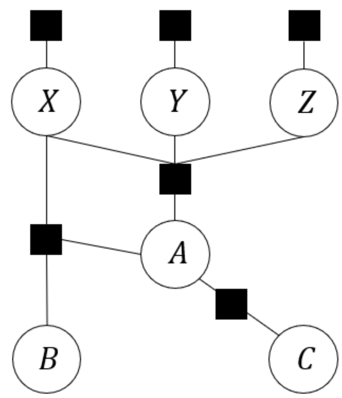


b) Draw and label the Factor Graph and Bayesian Network model which represents the following joint distribution $P(X_1, \ldots, X_5) = P(X_1)\, P(X_2)\, P(X_3|X_1, X_2)\, P(X_4|X_2, X_3) P(X_5|X_1, X_3)$.





4. For the factor graph given below, which of the following conditional independence relationship is true. Justify your answer.

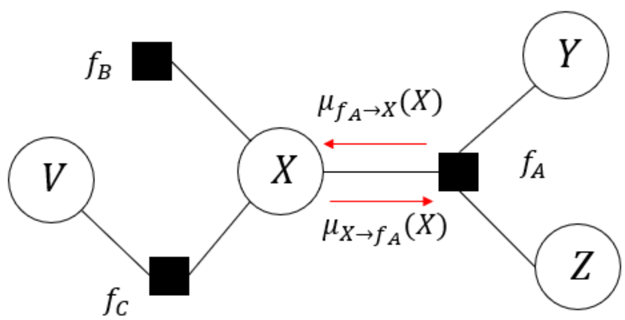


a) $C \perp\!\!\!\perp X \mid A$
Yes. Every path from X to C has node A. Therefore, if A is observed, then X and C are conditionally independent.

b) $B \perp\!\!\!\perp Y \mid X$
No. Even after X is observed, there is a direct path from B to Y via A.

c) $B \perp\!\!\!\perp Z \mid A, X$
Yes. If A, X are observed, then every path from B to Z has at least one observed node.

5. Belief Propagation equations: For the factor graph given below, write the equations for messages between f_A and X.



$$\mu_{X \to f_A}(X) = \mu_{f_B \to X}(X)\ \mu_{f_C \to X}(X)$$

$$\mu_{f_A \to X}(x) = \sum_{Y,Z} f_A(X, Y, Z) \times \mu_{Y \to f_A}(Y) \times \mu_{Z \to f_A}(Z)$$

6. Word problem on belief propagation

Clinicians at the Mayo Clinic built a factor graph model that captures the relationship between lung cancer and smoking habits using their expertise and survey data. The relationship between cancer and smoking is summarized in Table 3.

- X denotes whether a person has cancer or not.
- Y denotes whether a person smokes or not.

Table 1 Lung cancer statistics

X	f_1
No Cancer (0)	9
Cancer (1)	1

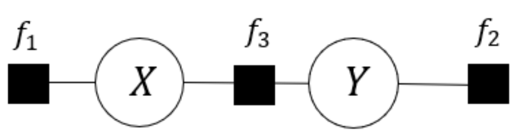
Table 2 Smoking Statistics

Y	f_2
Non-Smoker (0)	6
Smoker (1)	4

Table 3 Dependency between Lung Cancer and Smoking

f_3	$Y = Non\ Smoker$	$Y = Smoker$
$X = No\ Cancer$	8	3
$X = Cancer$	4	7

The factor graph model that captures information presented in Tables 1, 2 and 3 is:



Use the factor graph model you built above to answer the following questions.

a) Describe the joint distribution $P(X, Y)$ in terms of the factor functions. Remember to specify the equation for the normalization constant (partition function).

$$P(X, Y) = \frac{1}{Z} f_1(X) f_2(Y) f_3(X, Y)$$

$$Z = \sum_{X \in \{0,1\}} \sum_{Y \in \{0,1\}} f_1(X) f_2(Y) f_3(X, Y)$$

b) Compute the marginal probabilities by enumerating all combinations of X and Y.

X	Y	Product of factor functions	$P(X,Y)$
0	0	$f_1(0) \times f_2(0) \times f_3(0,0) = 432$	$\frac{1}{Z} \times f_1(0) \times f_2(0) \times f_3(0,0) = \frac{432}{592}$
0	1	$f_1(0) \times f_2(1) \times f_3(0,1) = 108$	$\frac{1}{Z} \times f_1(0) \times f_2(1) \times f_3(0,1) = \frac{108}{592}$
1	0	$f_1(1) \times f_2(0) \times f_3(1,0) = 24$	$\frac{1}{Z} \times f_1(1) \times f_2(0) \times f_3(1,0) = \frac{24}{592}$
1	1	$f_1(1) \times f_2(1) \times f_3(1,1) = 28$	$\frac{1}{Z} \times f_1(1) \times f_2(1) \times f_3(1,1) = \frac{28}{592}$
Total		$Z = 592$	1

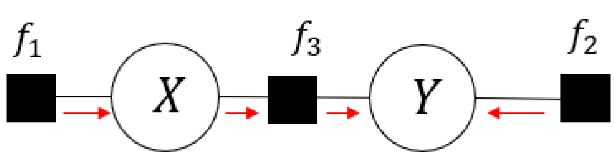
$P(X=0) = P(X=0, Y=0) + P(X=0, Y=1) = \frac{432+108}{592} = \frac{540}{592} = 0.91$

$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) = \frac{52}{592} = 0.09$

$P(Y=0) = P(X=0, Y=0) + P(X=1, Y=0) = \frac{456}{592} = 0.77$

$P(Y=1) = P(X=0, Y=1) + P(X=1, Y=1) = \frac{136}{592} = 0.23$

c) Computing marginal probability $P(Y)$ using belief propagation



i. First, calculate the message from f_3 to Y.

$$\mu_{f_1 \to X}(X) = f_1(X) = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$$

$$\mu_{X \to f_3}(X) = \mu_{f_1 \to X}(X) = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$$

Equation for message from f_3 to Y: $\mu_{f_3 \to Y}(Y) = \sum_X f_3(X,Y)\, \mu_{X \to f_3}(X)$

$$\mu_{f_3 \to Y}(Y=0) = f_3(0,0)\mu_{X \to f_3}(0) + f_3(1,0)\mu_{X \to f_3}(1) = 8 \times 9 + 4 \times 1 = 76$$

$$\mu_{f_3 \to Y}(Y=1) = f_3(0,1)\mu_{X \to f_3}(0) + f_3(1,1)\mu_{X \to f_3}(1) = 3 \times 9 + 7 \times 1 = 34$$

ii. Next, calculate the message from f_2 to Y.

$$\mu_{f_2 \to Y}(Y) = f_2(Y) = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

iii. Calculate $P(Y)$.

Equation for marginal probability: $P(Y) \propto \mu_{f_3 \to Y}(Y)\mu_{f_2 \to Y}(Y)$

$$\mu_{f_3 \to Y}(0)\mu_{f_2 \to Y}(0) = 76 \times 6 = 456$$

$$\mu_{f_3 \to Y}(1)\mu_{f_2 \to Y}(1) = 34 \times 4 = 136$$

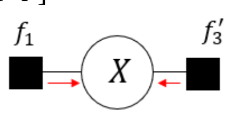
Exact computation taking into account the normalization:

$$P(Y=0) = \frac{456}{456+136} = \frac{456}{592} = 0.77$$

$$P(Y=1) = \frac{136}{456+136} = \frac{136}{592} = 0.23$$

iv. Do your answers to parts b) and c) match for $P(Y)$?
Yes.

v. Calculate $P(X = cancer \mid Y = Smoker)$ using Belief Propagation. Compare the risk of cancer on smoking with the probability of cancer in the overall population? [Hint: Modify the factor graph to account for the observation of Y]



$$f_3'(X) = f_3(X, Y=1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

$$\mu_{f_1 \to X}(X) = f_1(X) = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$$

$$\mu_{f_3' \to X}(X) = f_3'(X) = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$$

$$P(X|Y=1) \propto \mu_{f_1 \to X}(X)\mu_{f_3' \to X}(X) = \begin{bmatrix} 9 \times 3 \\ 1 \times 7 \end{bmatrix} = \begin{bmatrix} 27 \\ 7 \end{bmatrix}$$

$$P(X=1|Y=1) = \frac{7}{27+7} = \frac{7}{34} = 0.21$$

Probability of cancer increased from 0.09 ($P(X=1)$) to 0.21 ($P(X=1|Y=1)$).