ECE/CS 498 DSU/DSG Spring 2019
In-Class Activity 2

NetID: ____

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

(i) Review how to go from the word description of a problem to Bayesian Network
(ii) Factorize the joint distribution using conditional independence assumptions depicted by the BN
(iii) Performing inference using the factorized joint distribution

Problem: Our TA Chang can't start her day without coffee. So, she walks up to the coffee machine in her office. The coffee machine is unique. It can be either be working properly or have become faulty. Consequently, the taste of the coffee it produces can be good or bad depending on whether the machine was faulty or not. The taste of the coffee also depends on the amount of coffee powder in the machine, and the amount of milk in the machine. Those two quantities are maintained independent from the state of the machine and one another. There is an LED on the machine that glows to indicate that the machine is working properly. However, the LED itself is not very reliable, so sometimes it can indicate the wrong machine state.

Today, the LED is glowing. The amount of coffee in the machine is high (can be high, medium or low), and the amount of milk is high (can be high or low). Chang gets her coffee, but it tastes so bad that she is forced to spits it out. It tastes bad! Chang wants to figure out if the coffee machine has gone faulty. Can you help her?

1. What are the variables in the above problem? What values can the variables take?
   Coffee machine (V) = {Working, Faulty}
   Coffee powder (C) = {low, medium, high}
   Milk (M) = {low, high}
   Taste of the coffee (T) = {good, bad}
   LED glow (L) = {ON, OFF}

2. How many parameters would a full joint distribution over the above variables require?
   Total number of parameters = product of number of values each parameter can take – 1
   = 2x3x2x2x2 = 48 – 1 = 47

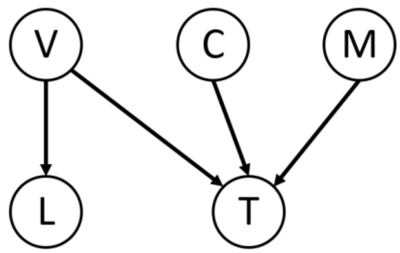
The number of parameters is high! You recall that you recently learned in class that Bayesian Networks can reduce the number of parameters. You think to yourself – maybe I can use a Bayesian Network (BN) to help Chang.

3. What are the nodes of the BN for the above problem?
   The variables defined in question 1. V, C, M, T, L.

4. What are the edges of the BN for the above problem? Describe them based on the problem description.
   Coffee machine (V) effects LED glow (L)

Coffee machine (V) effects Taste (T)
Amount of coffee powder (C) effects Taste (T)
Amount of milk (M) effect Taste (T)

5. Draw the BN based on your answer for Question 3 and 4.



Graph constructed! You look up your notes to see the next steps. The lecture slides state that factorizing the joint distribution requires applying the chain rule and then applying local semantics (conditional independence assumptions) as specified by the structure of the graph.

6. Apply chain rule to the joint distribution. [Hint: Recall that the parents must be conditioned on.]
   P(L, T, C, M, V) = P(L|T, C, M, V) P(T|C, M, V) P(C|M, V) P(M|V) P(V)

7. Simplify the terms of the joint distribution by using conditional independence assumptions. Specify the factorized joint distribution.
   P(L|T, C, M, V) = P(L|V)
   P(T|C, M, V) = P(T | C, M, V)
   P(C|M, V) = P(C)
   P(M|V) = P(M)

   P(L, T, C, M, V) = P(L|V) P(T|C, M, V) P(C) P(M) P(V)

8. How many parameters are required to specify the factorized joint distribution?
   P(V): 1, P(C): 2, P(M): 1, P(T| C, M, V): 12, P(L|V): 2
   Total = 1 + 2 + 1 + 12 + 2 = 18

Good going! You have constructed the BN, factorized the joint distribution and are satisfied that it needs fewer parameters. However, there are steps to complete to help Chang. The lecture slides suggest you need CPTs. Chang does not have training data for calculating CPTs. You look up the coffee machine online and found that the datasheet manual of coffee machine provided by its manufacturers has the information you need (the company is Bayesian Network savvy). Below are the CPTs.

V	P(L=ON|V)
Working	0.9
Faulty	0.2

V	C	M	P(T=good|C,M,V)
Working	Low	Low	0.5
Working	Low	High	0.6
Working	Medium	Low	0.6
Working	Medium	High	0.7
Working	High	Low	0.8
Working	High	High	0.9
Faulty	Low	Low	0.1
Faulty	Low	High	0.2
Faulty	Medium	Low	0.3
Faulty	Medium	High	0.35
Faulty	High	Low	0.35
Faulty	High	High	0.4

P(V=working) = 0.9

P(C=low) = 0.2
P(C=medium) = 0.6

P(M=low) = 0.5

9. Express the question about the machine being faulty given all other parameters as a probability expression.
Is P(V=faulty | T=bad, L=ON, C=high, M=high) > P(V=working | T=bad, L=ON, C=high, M=high)?

10. Simplify the expression from Question 9 using the factorized joint distribution, Bayes theorem and the theorem of total probability.

$$P(V|T,L,C,M) = \frac{P(V,T,L,C,M)}{P(T,L,C,M)}$$
$$= \frac{P(L|V)P(T|C,M,V)P(C)P(M)P(V)}{P(T,L,C,M)}$$

11. Is the coffee machine faulty given the observation? [Hint: Substitute the value from CPTs in the expression in 10 to evaluate your final answer. Remember that you might want to apply MAP rule here.]
Since we want to infer which probability is higher, and the denominator is just a scaling factor, we ignore it and only compute the numerator.

$$P(V = faulty|T = bad, L = ON, C = high, M = high)$$
$$\propto P(L = ON|V = faulty)P(T = bad|C = high, M = high, V = faulty)$$
$$P(C = high)P(M = high)P(V = faulty)$$
$$= 0.2 * 0.6 * 0.2 * 0.5 * 0.1 = 120 \times 10^{-5}$$

$$P(V = working|T = bad, L = ON, C = high, M = high)$$
$$\propto P(L = ON|V = working)P(T = bad|C = high, M = high, V = working)$$
$$P(C = high)P(M = high)P(V = working)$$
$$= 0.9 * 0.1 * 0.2 * 0.5 * 0.9 = \mathbf{810 \times 10^{-5}}$$

By applying MAP rule, we get that the machine is working. (She just had a bad day. :P)

Show your work to Chang. You have helped her and can feel happy about it. ☺