Assignment 4
Probabilistic Models of Human and Machine Intelligence
CSCI 7222

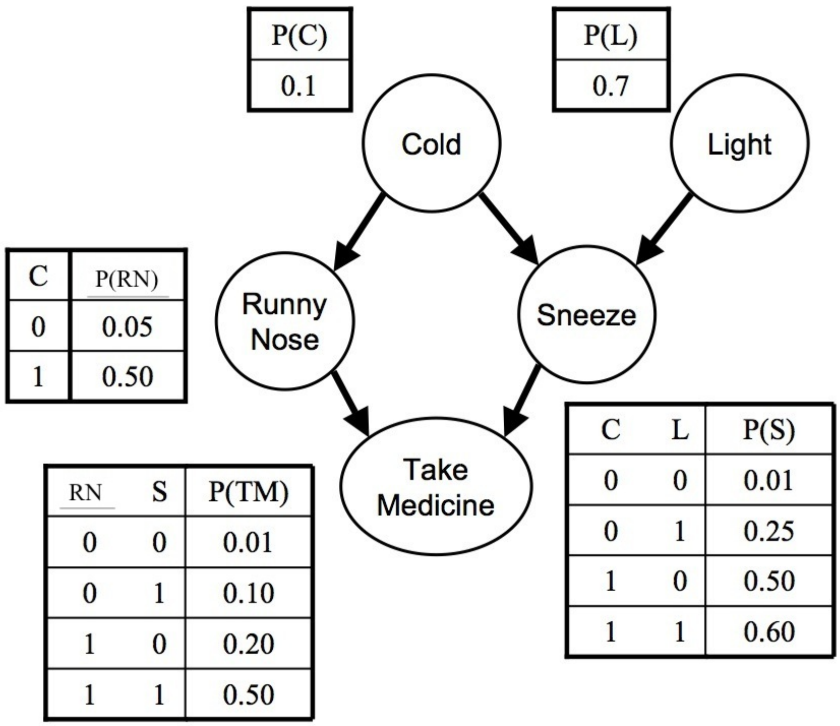
Assigned 9/23/2015
Due 10/6/2015

Goal

This assignment will give you practice in determining conditional independence and performing exact inference.

PART I

Consider the Bayes net with 5 binary random variables. (Thank you to Pedro Domingos of U. Washington for the figure.)



P(C)
0.1

P(L)
0.7

C	P(RN)
0	0.05
1	0.50

C	L	P(S)
0	0	0.01
0	1	0.25
1	0	0.50
1	1	0.60

RN	S	P(TM)
0	0	0.01
0	1	0.10
1	0	0.20
1	1	0.50

1. What is the Markov blankot of Sneeze?

2. What is the Markov blanket of Take Medicine?

3. Write the expression for the joint probability, P(L, C, RN, S, TM), in terms of the conditional probability distributions.

4. Using your formula in part 3, compute P(C=1 | TM=1, RN=0, L=0). Remember the definition of conditional probability: P(X | Y) = P(X, Y) / P(Y). Y can refer to multiple random variables.

5. Moralize the above graph to obtain an equivalent Markov net.

6. Write the joint probability function for the Markov net, with one term per maximal clique. Show the correspondence between the potential functions in this equation with the conditional probability distributions in question 3.

7. Is the Bayes net above a polytree? If not, what links might you add or remove to make it into a polytree? (A polytree is a directed graph with no undirected cycles; it is a generalization of a tree that allows for a node to have multiple parents.)

7. Which of the following are true:
(a) C ? TM | RN, S
(b) TM ? C | S

(c) C ? L
(d) C ? L | TM
(e) RN ? L | TM
(f) RN ? L
(g) RN ? L | S
(h) RN ? L | C, S
As explained in class, the expression "X ? Y | Z" means X and Y are independent when conditioned on Z.

PART II

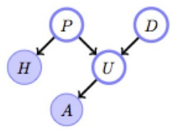


Figure 3.14: Party animal. Here all variables are binary. P = Been to Party, H = Got a Headache, D = Demotivated at work, U = Underperform at work, A =Boss Angry.

1. Write the joint distribution over A, D, U, H, and P in this graphical model (from Barber, Fig 3.14). Ignore the shading of H and A.

2. Write out an expression for P(H) as a summation over nuisance variables in a manner that would be appropriate for efficient variable elimination. (Don't do any numerical computation; simply write the expression in terms of the conditional probabilities and summations over the nuisance variables. Arrange the summations as we did in the variable elimination examples to be efficient in computing partial results.)

3. Write out a simplified expression for P(U=u | D=d) in a manner that would be appropriate for variable elimination. By 'simplified expression', I mean to reduce the expression to the simplest possible form, eliminating constants and unnecessary terms.

PART III

The Bayes net below comes from Kevin Murphy's text (Exercise 20.3).

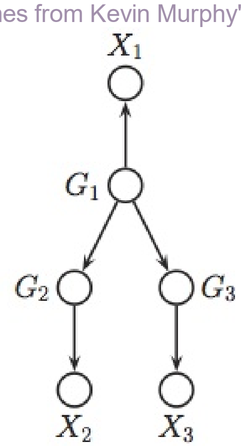


Figure 20.10 A simple DAG representing inherited diseases.

Consider the DGM in Figure 20.10 which represents the following fictitious biological model. Each G represents the genotype of a person: $G_i = 1$ if they have a healthy gene and $G_i = 2$ if they have a unhealthy gene. G_2 and G_3 may inherit the unhealthy gene from their parent G_1. $X_i \in \mathbb{R}$ is a continuous measure of blood pressure, which is low if you are healthy and high if you are unhealthy. We define the CPDs as follows

$$p(G_1) = [0.5, 0.5] \quad (20.7$$

$$p(G_2|G_1) = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \quad (20.7$$

$$p(G_3|G_1) = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \quad (20.7$$

$$p(X_i|G_i = 1) = \mathcal{N}(X_i|\mu = 50, \sigma^2 = 10) \quad (20.7$$

$$p(X_i|G_i = 2) = \mathcal{N}(X_i|\mu = 60, \sigma^2 = 10) \quad (20.7$$

The meaning of the matrix for $p(G_2|G_1)$ is that $p(G_2 = 1|G_1 = 1) = 0.9$, $p(G_2 = 1|G_1 = 2) = 0.$ etc.

1. Compute $P(G_1 = 2 \mid X_2 = 50)$

2. Compute $P(X_3 = 50 \mid X_2 = 50)$