MACHINE LEARNING COMS 4771, HOMEWORK 5
Assigned November 23, Due December 9, 2014 before 1:10pm.

Problem 1 (10 points)

Assume you are a contestant of a game show in which you are presented with three closed doors A, B, and C. Behind one of the doors is a car which will be yours if you choose the right door. After you have randomly (as you have no prior information) selected a door (say door A), the game host opens door B which has nothing inside, while keeping door A and C closed. The host then asks whether you want to change your selection from A to C. Should you change? Explain your answer.

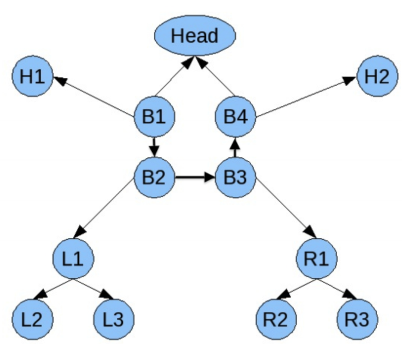
Problem 2 (15 points)

Let W, X, Y and Z be random variables, show the following:

- Prove that $X \perp Y|Z$ if and only if the joint probability $P(x,y,z)$ can be expressed in form of $a(x,z)b(y,z)$.
- Prove(or disprove with a counter-example) that $X \perp Y|Z$ and $X \perp W|(Y,Z)$ imply $X \perp (W,Y)|Z$
- Prove(or disprove with a counter-example) that $X \perp Y|(Z,W)$ and $X \perp Z|(Y,W)$ imply $X \perp (Y,Z)|W$
- **BONUS (5 points)** Given directed acyclic graphical model composed of nodes $X_1, X_2, \cdots, X_n$, prove that the conditional independent statements $X_i \perp X_{\text{Non-descendent of } i \setminus \pi_i} | X_{\pi_i}$ imply that the joint distribution can be decomposed as follows:

$$P(X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} P(X_i | X_{\pi_i})$$

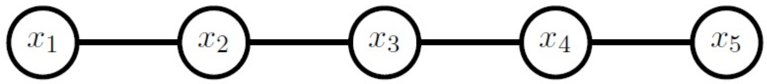
Problem 3 (5 points)



Eve is looking for WallE using her cameras but can't find WallE. Eve has small circuits for performing the sum-product junction-tree algorithm. Help her out by helping in building a WallE classifier! Design a junction-tree from the graph above which Eve has in her mind for WallE.

Problem 4 (20 points)

Consider the family of undirected graphical models known as Markov chains as shown below.



For simplicity, assume all variables in the model are binary. The probability distribution implied by the undirected graph is $p(x_1, ..., x_5) = \frac{1}{Z}\psi(x_1, x_2)\psi(x_2, x_3)\psi(x_3, x_4)\psi(x_4, x_5)$. Write an implementation of the junction tree algorithm that computes all the pairwise marginals $p(x_i, x_{i+1})$ for such a Markov chain for any number of variables n and any initialization of the clique potential functions. The initial clique potentials will serve as inputs to your junction tree algorithm and should be a cell array of $n-1$ elements, each of which is a 2 x 2 matrix of non-negative values as shown in the following Matlab code:

```
n = 5;
psis = cell(n-1, 1);
for i = 1:(n-1)
  psis{i} = rand(2,2);
end
```

The output of your junction tree algorithm (JTA) should be an identical data structure. It should contain consistent marginals that sum to unity appropriately and agree pairwise. Since the tree is only a chain, you don't have to implement a recursive algorithm (i.e. the Collect and Distribute steps in the Jordan book). Instead, you only need to perform left to right message passing and then right to left message passing by using a for loop or standard iteration. In other words, the JTA should process the cliques for $i = 1 : n-1$ and then for $i = n-1 : -1 : 1$ to do all the necessary messages in the JTA.

Test your algorithm by recovering the pairwise marginals when given the following values for the potential functions in the graphical model. Show code, results, and discussion.

$$\psi(x_1, x_2) = \left\{ \begin{array}{c|cc} & x_2 = 0 & x_2 = 1 \\ \hline x_1 = 0 & 0.1 & 0.7 \\ x_1 = 1 & 0.8 & 0.3 \end{array} \right\}$$

$$\psi(x_2, x_3) = \left\{ \begin{array}{c|cc} & x_3 = 0 & x_3 = 1 \\ \hline x_2 = 0 & 0.5 & 0.1 \\ x_2 = 1 & 0.1 & 0.5 \end{array} \right\}$$

$$\psi(x_3, x_4) = \left\{ \begin{array}{c|cc} & x_4 = 0 & x_4 = 1 \\ \hline x_3 = 0 & 0.1 & 0.5 \\ x_3 = 1 & 0.5 & 0.1 \end{array} \right\}$$

$$\psi(x_4, x_5) = \left\{ \begin{array}{c|cc} & x_5 = 0 & x_5 = 1 \\ \hline x_4 = 0 & 0.9 & 0.3 \\ x_4 = 1 & 0.1 & 0.3 \end{array} \right\}$$