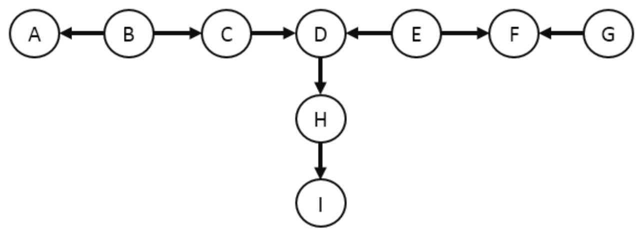
EE 583 Probabilistic Graphical Models

Homework 2

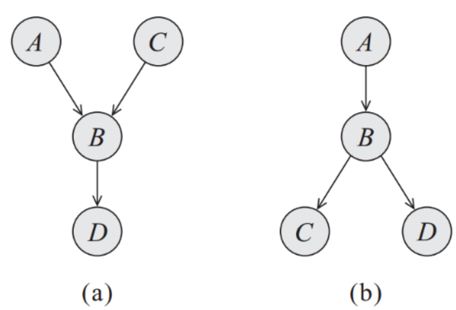
Due on Friday, March 25, 2016

1. **[9 pts]** According to the Bayesian Network shown below, list the cases, in which variables A and G become independent.



2. **[20 pts]** Let G be a BN structure and P be a distribution over random variables $\boldsymbol{X} = \{X_1, \dots, X_n\}$. Show that if P factorizes according to G, then $(X_i \perp \boldsymbol{X}_{\text{NonDes}(i)} | \boldsymbol{X}_{\text{Pa}(i)}) \in \text{I}(P)$ for all X_i.
Hint: Write $P(\boldsymbol{X}) = P(X_i, \boldsymbol{X}_{\text{Pa}(i)}, \boldsymbol{X}_{\text{NonDes}(i)}, \boldsymbol{X}_{\text{Des}(i)})$ and using factorization of the RHS, show $P(X_i | \boldsymbol{X}_{\text{Pa}(i)}, \boldsymbol{X}_{\text{NonDes}(i)}) = P(X_i | \boldsymbol{X}_{\text{Pa}(i)})$.

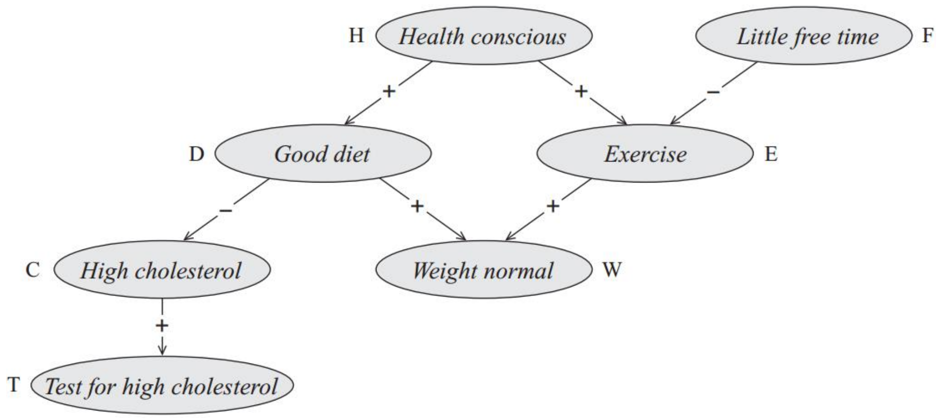
3. **[8 pts]** Consider the two networks.



For each of them, determine whether there can be any other Bayesian network that is I-equivalent to it.

4. **[20 pts]** Show that global Markov property of Bayesian networks imply the local Markov property. In other words, show that in a Bayesian network, each variable is d-separated from its non-descendants by its parents.

5. [18 pts] Consider the Bayesian network shown below. Assume all variables are binary valued. Suppose the signs at edges indicate how each random variable affects its child qualitatively, that is, $X \xrightarrow{+} Y$ indicates $P(Y=1|X=1, \boldsymbol{U}=\boldsymbol{u}) > P(Y=1|X=0, \boldsymbol{U}=\boldsymbol{u})$ for all values of $\boldsymbol{u}$ of Y's other parents (accordingly, the inequality reverses when $X \xrightarrow{-} Y$).



For each pair of the probabilities listed in the rows of the table below, identify which one is larger than the other, or if they are equal, or if they are incomparable with the given edge-sign information alone.

(a)	$P(T=1\|D=1)$	$P(T=1)$
(b)	$P(D=1\|T=0)$	$P(D=1)$
(c)	$P(H=1\|E=1, F=1)$	$P(H=1\|E=1)$
(d)	$P(C=1\|F=0)$	$P(C=1)$
(e)	$P(C=1\|H=0)$	$P(C=1)$
(f)	$P(C=1\|H=0, F=0)$	$P(C=1\|H=0)$
(g)	$P(D=1\|H=1, E=0)$	$P(D=1\|H=1)$
(h)	$P(D=1\|E=1, F=0, W=1)$	$P(D=1\|E=1, F=0)$
(i)	$P(T=1\|W=1, F=0)$	$P(T=1\|W=1)$

6. [25 pts] Download the `diningData.mat` and `categoryNames.mat` files from Moodle. Each column of the `diningData` matrix corresponds to one real dining scene image from JHU table setting dataset (http://cis.jhu.edu/entropy-pursuit/images.php), and each row corresponds to a particular object category from a refined set of $n = 10$ categories commonly seen on table settings. The (i, j)th entry of `diningData` is the binary presence indicator of category i in image j. Similarly, the ith entry of `categoryNames` array contains the name for the ith category.

a) Let $X_i \in \{0,1\}$ denote the indicator random variable for $i = 1, \dots, n$. Using data, estimate and provide the mutual information $\hat{I}(X_i; X_j)$ for each pair $i \neq j$.

b) Setting $\hat{I}(X_i; X_j)$ as edge weights between each possible variable pair, write a program to compute the maximally spanning tree T^* over $\boldsymbol{X} = (X_i, i = 1, \dots, n)$, and plot (or draw) T^* providing category names as node labels.

c) Let $P(\boldsymbol{X})$ be the true joint distribution over $\boldsymbol{X}$, and let $P_T(\boldsymbol{X})$ denote some approximation to P that uses tree structure T. Show that $T^* = \arg\max_{T} D_{\text{KL}}(P||P_T)$.