**CPSC-340** **Machine Learning and Data Mining** **2012**

**Homework # 2 - Due Friday 28th of September**

NAME:_____

Signature:_____

STD. NUM: _____

1. **Inference in the sprinkler network**

   Please calculate the two posteriors, $P(S = 1|W = 1)$ and $P(S = 1|W = 1, R = 1)$ in the sprinkler network described in class. (Show your calculations as well as giving the final result.)

   $P(S = 1|W = 1) =$

$$P(S = 1|W = 1, R = 1) =$$

2. **Conditional independence**

(a) Let $H \in \{1, \ldots, K\}$ be a discrete random variable that can take $K$ values. Let $e_1$ and $e_2$ be the observed values of two other random variables $E_1$ and $E_2$. Suppose we wish to calculate the vector

$$\vec{P}(H|e_1, e_2) = (P(H = 1|e_1, e_2), \ldots, P(H = K|e_1, e_2))$$

Which of the following sets of numbers are sufficient for the calculation?

i. $P(e_1, e_2)$, $P(H)$, $P(e_1|H)$, $P(e_2|H)$

ii. $P(e_1, e_2)$, $P(H)$, $P(e_1, e_2|H)$

iii. $P(e_1|H)$, $P(e_2|H)$, $P(H)$

(b) Now suppose we now assume $(E_1 \perp E_2)|H$ (i.e., $E_1$ and $E_2$ are conditionally independent given $H$). Which of the above 3 sets are sufficent now?

Show your calculations as well as giving the final result. Hint: use Bayes rule.

3. **Probability and Bayes Rule**

A new band called The XX (aka $B_2$) is inspired by an old band called Biggie and Tupac (aka $B_1$). 90% of music critics think Biggie and Tupac was a great ($G$) band, 10% that it was moderate ($M$) and 0% that it was awful ($A$). These critics have also compiled the following table:

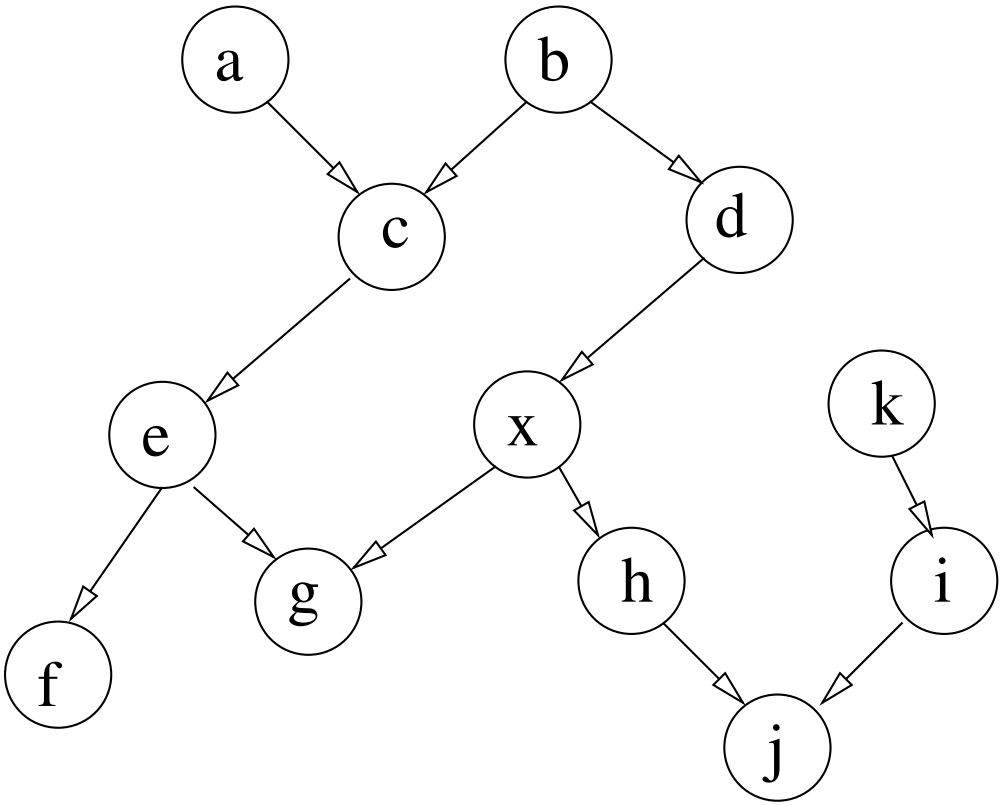
$$\begin{array}{cc} & \begin{array}{ccc} & B_2 & \\ G & M & A \end{array} \\ \begin{array}{c} \\ B_1 \\ \\ \end{array} & \begin{array}{c} G \\ M \\ A \end{array} \begin{pmatrix} 1 & 0 & 0 \\ 0.1 & 0.9 & 0 \\ 0.2 & 0.3 & 0.5 \end{pmatrix} \end{array}$$

The table says that the probability of a new band ($B_2$) being great given that the inspiring band ($B_1$) was great is $P(B_2 = G|B_1 = G) = 1$. Similarly, $P(B_2 = G|B_1 = M) = 0.1$, $P(B_2 = M|B_1 = A) = 0.3$, and so on. Note that the numbers in the rows add up to 1, so the table is a probability transition matrix.

(a) Given what the critics think of Biggie and Tupac and the fact that the Biggie and Tupac inspired The XX, what is the probability that The XX is a great band? i.e. $P(B_2 = G) =?$

(b) What is $P(B_1 = G|B_2 = G)$ in this example?

4. Given the following directed acyclic probabilistic graphical model:




(a) What nodes in the above picture correspond to the Markov blanket of node $x$.

(b) Is $i$ independent of $k$ given $j$. Answer yes or no.

(c) Is $b$ independent of $a$ given $c$ and $e$ and $d$. Answer yes or no.

5. **Probability** Let $P(HIV = 1) = 1/500$ be the probability of contracting HIV by having unprotected sex once. If one has unprotected sex twice, what is the probability of contracting HIV? What if we have unprotected sex 500 times?