EE512 Graphical Models | **University of Washington**
Spring 2006 | **Dept. of Electrical Engineering**

Homework 1: Due Friday, April 7, 2006, 5:00pm

Prof: Jeff A. Bilmes `<bilmes@ee.washington.edu>` TA: Chris Bartels `<bartels@ee.washington.edu>`

Problem 1. Factorization and BNs.

a) Consider the following probability distribution over 6 variables A,B,C,D,E, and F for which the factorization as stated below holds. Find and draw a Bayesian network that for which this factorization is true, but for which no additional factorizations nor any fewer factorizations are true.

$$p(a, b, c, d, e, f) = p(a)p(b)p(c|a, b)p(d|b)p(e|c, d)p(f|e)$$

1b) Given your BN from problem 1, state whether the following independence statements are true or false. Indicate the justification used in each case.

$b \perp\!\!\!\perp e | c$

$a \perp\!\!\!\perp d | e$

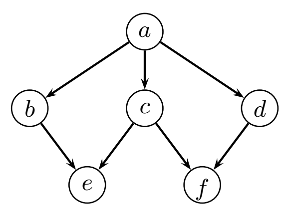
$a \perp\!\!\!\perp d | f$

$c \perp\!\!\!\perp d | b$

.

Problem 2. Independence and algebra.

Consider the following BN as shown. Show, using only algebraic equations that $b \perp\!\!\!\perp f | a$ is true.



Problem 3. Independence and Conditional Independence

Given a set V of N random variables, and where we have $A, B, C \subseteq V$ and A, B, C are disjoint.

3a) Describe a set of random variables X_A and X_B and X_C such that $X_A \perp\!\!\!\perp X_B | X_C$ but also $X_A \not\perp\!\!\!\perp X_B$.

3b) Describe a set of random variables X_A and X_B and X_C such that $X_A \not\perp\!\!\!\perp X_B | X_C$ but also $X_A \perp\!\!\!\perp X_B$.

Problem 4. Factorization and Independence

In class, we spoke about factorization and independence. We said that conditional independence always involves some factorization, but not vice versa. In this problem, your task is to come up with a probability distribution over at least 3 variables which involves some factorization constraint, but for which there are no independence assumptions (i.e., nothing is either independent nor conditionally independent of anything else).

Problem 5. Independence

We are given a set of N random variables $X_1, X_2, \ldots, X_N$. Prove that if all the variables are mutually independent (i.e., that $X_A \perp\!\!\!\perp X_B \;\forall A, B \subseteq V$), then all independence statements hold (i.e., $X_A \perp\!\!\!\perp X_B | X_C$ for any disjoint A, B, C such that $\emptyset \subseteq A \subseteq V$, $\emptyset \subseteq B \subseteq V$, $\emptyset \subseteq C \subseteq V$). Show also that if $X_1, X_2, \ldots, X_N$ are only pairwise independent (i.e., that $X_i \perp\!\!\!\perp X_j \;\forall i \neq j$) then it is not necessarily the case that all independence statements hold (i.e., give a counterexample).

Problem 6:

We defined the (DF) (directed factorization) property of Bayesian networks to be such that the BN over n nodes corresponds to all distributions over n variables (each of which is defined on a finite sized discrete alphabet) that factorize as follows:

$$p(X_{1:N}) = \prod_{i=1}^{n} f_i(x_i, x_{\pi_i})$$

Where π_i are the parents of i in the graph, and where $X_{1:N}$ corresponds to the set of N variables $\{X_1, X_2, \ldots, X_N\}$. We constrain the f_i functions so that $0 \leq f_i(x_i, x_{\pi_i}) \leq 1$ and $\sum_{x_i} f_i(x_i, x_{\pi_i}) = 1$. Your task is to prove, under these constraint, that $f_i(x_i, x_{\pi_i}) = p(x_i | x_{\pi_i})$.

Problem 7:

There are three parts to this question. Explain both (1) using words and (2) mathematically why is it not valid to have directed cycles in a Bayesian network. Also we described how a Bayesian network is sometimes given a causal interpretation, but how it is best to think of a BN as specifying a factorization rather than ascribing such a casual understanding. Even so, suppose a BN loosely represented causality in some limited circumstances. (3) Would any problems arise in this causal interpretation if a BN was allowed to have directed cycles?