Solutions: Homework 2
Belief Networks ICS 275b

Rina Dechter

Problem (1 3.1 Pearl's book)

$I(X, Y, Z) \& I(XY, Z, W) \rightarrow I(X, Y, Z) \& I(X, YZ, W)$ ....(Using weak union and symmetry)
$\rightarrow I(X, Y, ZW)$ ....(Using contraction).
$\rightarrow I(X, Y, W)$..................(Using decomposition).
This proves the chaining rule.

To prove that the chaining cannot replace equation (3.6d) in the set of axioms, we give an example in which apart from the chaining rule all the set of axioms except equation (3.6d) are satisfied. Let $U = \{X, Y, Z, W\}$ be the set of variables.

Let $I = \{\{X, Y, W\}, \{X, Y, Z\}, \{XY, Z, W\}, \{X, Z, W\}, \{Y, Z, W\}, \{X, ZY, W\}, \{Y, ZX, W\}$, *all symmetric relations* $\}$.

I satisfies symmetry, decomposition, chaining rule, intersection and weak union but it does not satisfy equation (3.6d) (contraction). This is because although $\{X, Y, Z\}$ and $\{X, ZY, W\}$ are in I, $\{X, Y, ZW\}$ is not in I.

Problem 2

The dependencies based on the D-separation criteria are given in the following table:

Relation	D-separated
$I(E, G\|\phi)$	Yes
$I(C, D\|\phi)$	No
$I(C, D\|G)$	No
$I(B, C\|A)$	Yes
$I(\{C, D\}, E\|\phi)$	Yes
$I(F, \{E, H\}\|A)$	Yes
$I(\{A, C\}, \{H, E\}\|D)$	Yes

Problem 3

The following set of independencies satisfies (a),(b),(c) and (d) and is minimal.

$$\{I(X, \phi, Y), I(Y, \phi, X), I(X, \{Y, W\}, Z), I(Z, \{Y, W\}, X)\}.$$

A common mistake was to add a set of independencies to satisfy the graphoid properties.

Problem 4

The dependencies based on the joint probability distribution are given in following table:

Relation	Independent
$I(X, W\|Y)$	No
$I(X, W\|Z)$	No
$I(X, W\|Y, Z)$	Yes
$I(Y, Z\|X, W)$	No

Problem 5 (Problem 3.3 a,b and d)

3.3 a: Show that the graph G is a minimal I-map of P.

First, we show that G is an *I-map* of P and then show that it is minimal. For this, we have to show that all vertex-separations in G correspond to independencies:

- $\langle x|yz|w\rangle$
- $\langle y|xw|z\rangle$

$\forall\ x, y, z, w$ such that $P(w, y, z) > 0$, $P(x, yz, w) = P(x|y, z)$. This implies that $I(x, yz, w)$ holds.

$\forall\ x, y, z, w$ such that $P(y, w, z) > 0$, $P(y, w, z) = P(y|x, w)$. This implies that $I(y, xw, z)$ holds.

To show that G is minimal we have to delete edges of G one by one and check if any new independencies are introduced.

For example, if we delete the edge between x and y, we have $\langle x|z|w\rangle$ but $P(x = 1|z = 1, w = 1) = 1$ and $P(x = 1|z = 1) = 1/2$ and so the independency $I(x, z, w)$ implied by G is not in P. We can repeat the same for other edges.

So, the graph G' obtained by deleting any edge from G is not an I-map of P and so G is a minimal I-map of P.

3.3b: Show that P cannot be expressed as a product of functions on the cliques of G.

We prove this using contradiction. Let us assume that the join distribution P(x,y,z,w) is given by the following functions on the edges:

$$P(x, y, z, w) = f_1(x, y) * f_2(y, z) * f_3(z, w) * f_4(x, w)$$

Because $P(x = 1, y = 1, z = 1, w = 1) > 0$, $f_1(x = 1, y = 1) \neq 0$, $f_2(y = 1, z = 1) \neq 0$, $f_3(z = 1, w = 1) \neq 0$ and $f_4(x = 1, w = 1) \neq 0$.

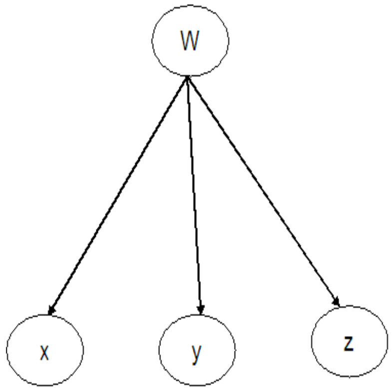
Because $P(x = 1, y = 2, z = 2, w = 2) > 0$, $f_1(x = 1, y = 2) \neq 0$, $f_2(y = 2, z = 2) \neq 0$, $f_3(z = 2, w = 2) \neq 0$ and $f_4(x = 1, w = 2) \neq 0$.

Because $P(x = 2, y = 2, z = 1, w = 3) > 0$, $f_1(x = 2, y = 2) \neq 0$, $f_2(y = 2, z = 1) \neq 0$, $f_3(z = 1, w = 3) \neq 0$ and $f_4(x = 2, w = 3) \neq 0$.

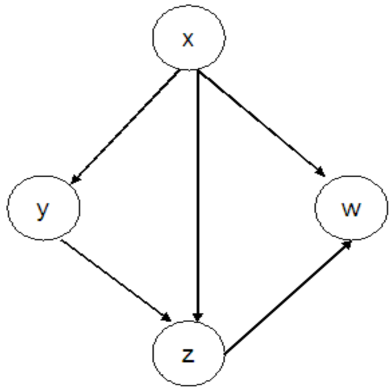
However, $P(x = 1, y = 2, z = 1, w = 1) = f_1(x = 1, y = 2) * f_2(y = 2, z = 1) * f_3(z = 1, w = 1) * f_4(x = 1, w = 1) > 0$, a contradiction.

3.3 d

The Bayesian Network for the ordering (w, x, y, z) is given below.



The Bayesian Network for the ordering (x, y, z, w) is given below.

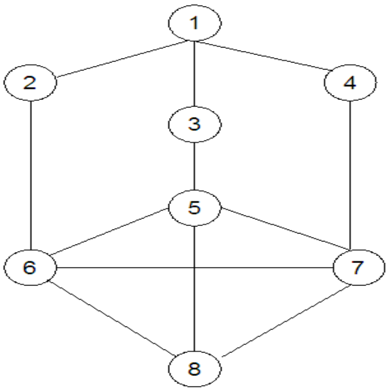


The parameters of these Bayesian Networks can be easily computed from the joint-distribution $P(x, y, z, w)$. For example, the distribution $P(x|w)$ can be computed as follows:

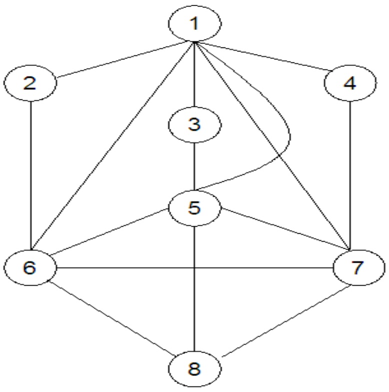
$$P(x|w) = \frac{\sum_{y,z} P(x,y,z,w)}{\sum_{x,y,z} P(x,y,z,w)}$$

Problem 7, Pearl's book 3.6 a,b,c,d

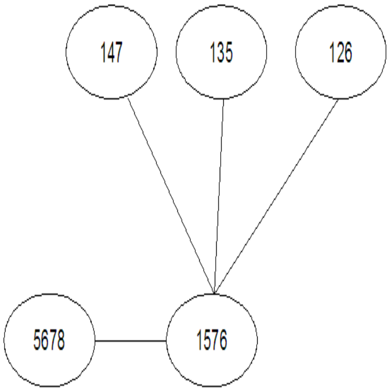
a The markov network is given below:



b Find an undirected graph G such that P is decomposable relative to G.



c Draw a join tree of G.



d Find an algebraic representation of P such that $P > 0$.

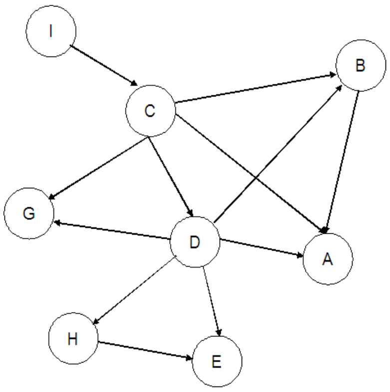
$$P(1,\ldots,8) = \frac{P(147)P(135)P(146)P(1567)P(5678)}{P(17)P(15)P(16)P(567)}$$

A common mistake[1] was to represent $P(1,\ldots,8)$ as follows:

$$P(1,\ldots,8) = P(1)P(2|1)P(3|1)P(4|1)P(5|3)P(6|2)P(7|4)P(8|567)$$

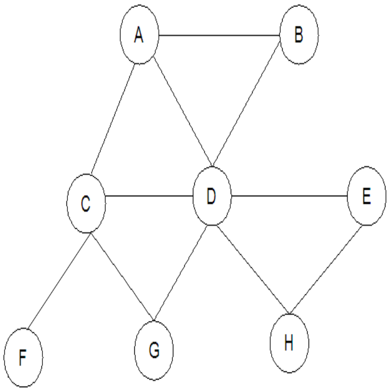
Problem 8

a The Bayesian Network along the ordering FCDBAHEG is given below:



[1]although it is correct based on theorem 10 in Pearl's book, I expected the answer given above

b The Markov Network is given below:



Problem 10

This problem was graded based on the structure and parameters of the submitted network and the constraints on probabilities. 5 points were taken off for each mistake in structure and each constraint violation.

A reference network is shown in the figure below. To be correct, a network did NOT have to exactly match this network because one could argue to delete or add a number of edges to the network.