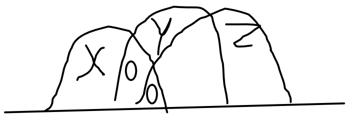
Assignment 4

Chapter 3

9. (a) Prove that under the constraint that $X \to Y \to Z$ forms a Markov chain, $X \perp Y|Z$ and $X \perp Z$ imply $X \perp Y$.

   (b) Prove that the implication in a) continues to be valid without the Markov chain constraint.

a)



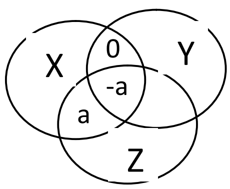
Since X->Y->Z forms a Markov chain, the information diagram is shown in the above diagram.

$X \perp Y | Z$ means I(X;Y|Z)=0

$X \perp Z$ means I(X; Z)=0

As shown above, these conditions imply that I(X;Y)=0, i.e., $X \perp Y$

b)



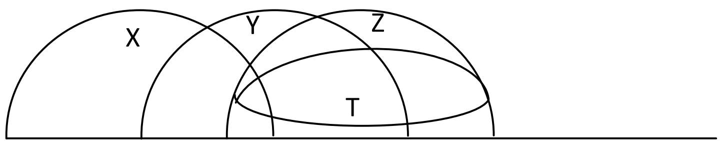
The conditions lead to the above information diagram. Since I(X;Y)>=0, a has to equal to 0. Thus $X \perp Y$.

10. (a) Show that $Y \perp Z|T$ does not imply $Y \perp Z|(X,T)$ by giving a counterexample.
(b) Prove that $Y \perp Z|T$ implies $Y \perp Z|(X,T)$ conditioning on $X \to Y \to Z \to T$.

a)

11. (a) Let $X \to Y \to (Z,T)$ form a Markov chain. Prove that $I(X;Z)+I(X;T) \leq I(X;Y)+I(Z;T)$.
(b) Let $X \to Y \to Z \to T$ form a Markov chain. Determine which of the following inequalities always hold:
i. $I(X;T)+I(Y;Z) \geq I(X;Z)+I(Y;T)$
ii. $I(X;T)+I(Y;Z) \geq I(X;Y)+I(Z;T)$
iii. $I(X;Y)+I(Z;T) \geq I(X;Z)+I(Y;T)$.

a)



From the above information diagram, we can see that

$$I(X;Z)+I(X;T)=I(X;ZT)+I(X;Z;T)\leq I(X;Y)+I(Z;T)$$

b)

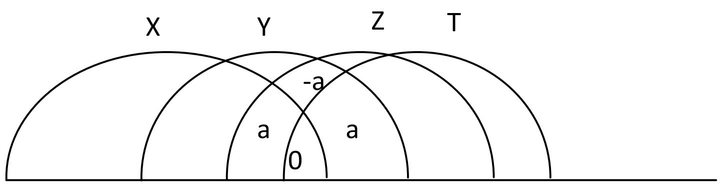
It can be seen from the information diagram that i) and iii) always hold.

14. Prove that for random variables X, Y, Z, and T,

$$\left.\begin{array}{l} X \perp Z|Y \\ (X,Y) \perp T|Z \\ Y \perp Z|T \\ Y \perp Z|X \\ X \perp T \end{array}\right\} \Rightarrow Y \perp Z.$$

Hint: Observe that $X \perp Z|Y$ and $(X,Y) \perp T|Z$ are equivalent to $X \to Y \to Z \to T$ and use an information diagram.

As given by the hint, X->Y->Z->T forms an Markov chain. Thus we should have the following information diagram.



Since u* is always non-negative, we have a=0. Thus I(Y;Z)=0 ⇔ $Y \perp Z$.

15. Prove that

$$\left.\begin{array}{l} X \perp Y \\ X \perp Y|(Z,T) \\ Z \perp T|X \\ Z \perp T|Y \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} Z \perp T \\ Z \perp T|(X,Y) \\ X \perp Y|Z \\ X \perp Y|T. \end{array}\right.$$

(Studený [346].)

Chapter 4

5. Prove directly that the codeword lengths of a prefix code satisfy the Kraft inequality without using Theorem 4.4.

A prefix code implies that the code only goes to leaves on a code tree. Since a full tree makes the Kraft inequality tight, all other cases should satisfy Kraft inequality.

7. *Suffix codes* A code is a suffix code if no codeword is a suffix of any other codeword. Show that a suffix code is uniquely decodable.

Just decode the suffix code stream backwards. Then suffix code turns into prefix code and is uniquely decodable.

9. *Random coding for prefix codes* Construct a binary prefix code with codeword lengths $l_1 \leq l_2 \leq \cdots \leq l_m$ as follows. For each $1 \leq k \leq m$, the codeword with length l_k is chosen independently from the set of all 2^{l_k} possible binary strings with length l_k according the uniform distribution. Let $P_m(good)$ be the probability that the code so constructed is a prefix code.

   (a) Prove that $P_2(good) = (1 - 2^{-l_1})^+$, where

$$(x)^+ = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

   (b) Prove by induction on m that

$$P_m(good) = \prod_{k=1}^{m} \left(1 - \sum_{j=1}^{k-1} s^{-l_j}\right)^+ .$$

   (c) Observe that there exists a prefix code with codeword lengths $l_1, l_2, \cdots, l_m$ if and only if $P_m(good) > 0$. Show that $P_m(good) > 0$ is equivalent to the Kraft inequality.

By using this random coding method, one can derive the Kraft inequality without knowing the inequality ahead of time. (Ye and Yeung [395].)

a) For the code with length l_2 to be prefix code, it should not be chosen as the child of the code with length l_1.

   The total number of choices of the code with length l_2 is 2^{l_2} .

   The number of children of the code with length l_1 that has the length of is l_2 $2^{l_2-l_1}$.

   So P2(good) = $\dfrac{2^{l_2} - 2^{l_2-l_1}}{2^{l_2}} = 1 - 2^{-l_1}$

b) As shown in a), the result holds for m=2.

Suppose the result holds for m. For m+1, the new code cannot be the child of any previous codes in a code tree. Thus

$$P_{m+1}(good) = P_m(good)\frac{2^{l_{m+1}} - 2^{l_{m+1}-l_1} - 2^{l_{m+1}-l_2} \ldots - 2^{l_{m+1}-l_m}}{2^{l_{m+1}}} = P_m(good)\left(1-\sum_{j=1}^{m} 2^{-l_j}\right).$$

If the new term in the product is less than zero, the probability is zero.

The result is proved.

c) Since we need Pm>0, the last term in the product needs to be greater than zero. This is equivalent to Kraft inequality.