2.6 Conditional mutual information vs. unconditional mutual information

Give examples of joint random variables X, Y , and Z such that  
(a) I (X; Y | Z) < I (X; Y).  
(b) I (X; Y | Z) > I (X; Y).

Solution:Conditional mutual information vs. unconditional mutual information.  
(a) The last corollary to Theorem 2.8.1 in the text states that if X -> Y -> Z that is, if p(x; y | z) = p(x j z)p(y j z) then, I(X; Y ) ≥ I(X; Y j Z) . Equality holds if and only if I(X; Z) = 0 or X and Z are independent.

A simple example of random variables satisfying the inequality conditions above is, X is a fair binary random variable and Y = X and Z = Y . In this case,

I(X; Y ) = H(X) - H(X | Y ) = H(X) = 1

and,

I(X; Y | Z) = H(X | Z) - H(X | Y, Z) = 0:

So that I(X; Y ) > I(X; Y | Z) .  
(b) This example is also given in the text. Let X; Y be independent fair binary random variables and let Z = X + Y . In this case we have that,

I(X; Y ) = 0

and,

I(X; Y | Z) = H(X | Z) = 1=2:

So I(X; Y ) < I(X; Y | Z) . Note that in this case X; Y; Z are not markov.

2.16 Bottleneck.

Suppose that a (nonstationary) Markov chain starts in one of n states, necks down to k < n states, and then fans back to m > k states. Thus, X1 → X2 → X3, that is, p(x1, x2, x3) = p(x1)p(x2|x1)p(x3|x2), for all x1 ∈ {1, 2, . . . , n}, x2 ∈ {1, 2, . . . , k}, x3 ∈ {1, 2, . . . , m}.  
(a) Show that the dependence of X1 and X3 is limited by the bottleneck by proving that I (X1; X3) ≤ log k.(b) Evaluate I (X1; X3) for k = 1, and conclude that no dependence can survive such a bottleneck.

**Solution:** Bottleneck.

1. From the data processing inequality, and the fact that entropy is maximum for a  
   uniform distribution, we get

I(X1; X3) ≤ I(X1; X2)

= H(X2) - H(X2 j X1)

≤ H(X2)  
≤ log k

Thus, the dependence between X1 and X3 is limited by the size of the bottleneck.  
That is I(X1; X3) ≤ log k .

1. For k = 1 , I(X1; X3) ≤ log 1 = 0 and since I(X1; X3) ≥ 0 , I(X1; X3) = 0 .

Thus, for k = 1 , X1 and X3 are independent.