School of Electrical and Computer Engineering
Georgia Institute of Technology
ECE 6605 Information Theory
Fall 2003

Homework #1
Assigned: August 20, 2003
Due: September 3, 2003

1) Example of joint entropy (entries in the following table are P(X=x,Y=y))

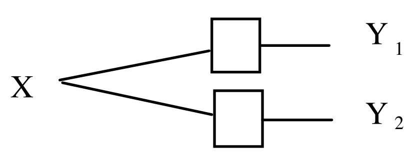
	Y=1	Y=2	Y=3
X=A	1/4	1/8	0
X=B	1/4	1/4	1/8

Find

(a) H(X), H(Y)
(b) H(X|Y), H(Y|X)
(c) H(X,Y)
(d) H(Y)-H(Y|X)
(e) I(X;Y)
(f) Draw a Venn diagram for the quantities in (a) through (e).

2) Mutual information examples
(a) Consider a fair coin flip. What is the mutual information between the top side and the bottom side of the coin?
(b) A 6-sided fair die is rolled. What is the mutual information between the top side and the front face (the side most facing you)?

3) Consider the following simple system



This problem is applicable to the situation where you get two independent (noisy) observations Y_1 and Y_2 of the random variable X (think of the boxes as communication channels). This is a simplified version of what happens in some multiple antenna communication system.

Let X, Y_1 and Y_2 be binary random variables. If $I(X; Y_1) = 0$ and $I(X; Y_2) = 0$,

does it follow that I(X; Y_1; Y_2) = 0?
a) Yes or no?
b) Prove or provide a counterexample.

4) <u>Processing on a random variable.</u> In many signal processing or communications applications an observation X is further processed by some operation g(X), e.g. filtering.

   Show that H(X) ≥ H(g(X)). Under what conditions on the function g() are they equal?

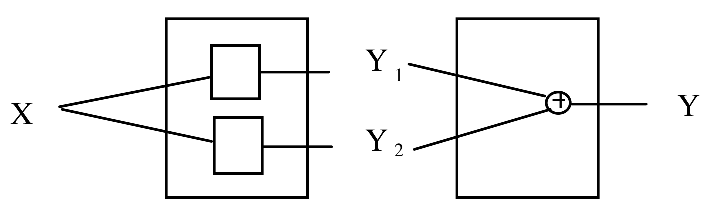
   Hints: 1) first think about what processing can do to the possible values of X and the resulting values g(X) and convince yourself this makes intuitive sense. 2) Form the joint entropy H(X,g(X)) and express it two different ways using the chain rule.

5) <u>Multiple observations-part 2.</u> In the figure of problem 3 assume that X is a random variable and

$$Y_1 = X + N_1$$
$$Y_2 = X + N_2$$

   where N_1 and N_2 are independent random variables with equal variance σ^2. The purpose of this problem is to see why multiple independent observations are better than a single observation.

   Show that a simple averaging receiver that performs the following operation $Y = (Y_1 + Y_2)/2$ has the effect of *<u>reducing the noise</u>* when compared with just a single noisy of X.



Channel Receiver