582650 Information-Theoretic Modeling (Fall 2014)

Homework 2 (due September 18)

1. Continuing with the *Alice in Wonderland* exercise (ex. 2) from last week, apply your ideas from last week to design a code for bigrams, i.e., combinations of two subsequent letters. In other words, encode the source (Chapter I) as `'ch' 'ap' 'te' 'r '`, ...

   Do you achieve better compression than last week?

2. Consider two random variables X and Y with the following joint distribution:

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	1/5	1/5	1/5
$X = 1$	0	1/5	1/5.

   In other words, the case $X = 1, Y = 0$ never occurs, but the other combinations are all equally probable.

   Calculate the values of

   (a) $H(X)$ and $H(Y)$

   (b) $H(X \mid Y)$ and $H(Y \mid X)$

   (c) $H(X, Y)$

   (d) $H(Y) - H(Y \mid X)$

   (e) $I(X; Y)$.

   Optionally, you may wish to draw a graphical representation to illustrate these quantities (as on page 13 of Lecture 3).

3. *(Exercise 2.14 in Cover & Thomas)* Let X and Y be random variables that take on values $x_1, \ldots, x_r$ and $y_1, \ldots, y_s$, respectively. Let $Z = X + Y$.

   (a) Show that $H(Z|X) = H(Y|X)$. Argue that if X, Y are independent, then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$. Thus, the addition of *independent* random variables adds uncertainty.

   (b) Give an example of (dependent) random variables in which $H(X) > H(Z)$ and $H(Y) > H(Z)$.

   (c) Under what conditions does $H(Z) = H(X) + H(Y)$?

4. Consider the simple Bernoulli model that generates independent random bits with $\Pr[X_i = 1] = p$ for some fixed $0 \leq p \leq 1$.

   For sequence length n, and some $\epsilon > 0$, the typical set A_ϵ^n is defined as the set of sequences $x_1, \ldots, x_n$ such that

$$2^{-n(H(X)+\epsilon)} \leq p(x_1, \ldots, x_n) \leq 2^{-n(H(X)-\epsilon)} \quad .$$

   (a) What are the sequences in the typical set $A_{0.1}^{100}$ under the Bernoulli model when $p = 0.1$?

   (b) What is the the probability of the typical set?

   (c) Compare this to the optimal lottery strategy from last week's exercises.

5. Let X and Y be two discrete random variables about which we only know that $H(X) = 3$ and $H(Y) = 4$. Under this assumption, what are the minimum and maximum values of the mutual information $I(X; Y)$? Justify your claims, and show by example that your upper and lower bounds can actually be achieved.