Homework 1
Due March 2, 2006

Prof. Amit Chakrabarti
Computer Science Department
Dartmouth College

Please typeset your solutions using LaTeX and submit a printed copy.

1. Let X be a discrete random variable. Show that the entropy of a function of X does not exceed that of X by justifying the following steps:

$$\begin{aligned} H(X, g(X)) &= H(X) + H(g(X) \mid X) && (1) \\ &= H(X)\,; && (2) \\ H(X, g(X)) &= H(g(X)) + H(X \mid g(X)) && (3) \\ &\geq H(g(X))\,. && (4) \end{aligned}$$

   Justify each of the equations (1), (2), (3) and (4).

2. Show that if $H(Y|X) = 0$, then Y is a function of X, i.e., for all x with $p(x) > 0$, there is only one possible value of y with $p(x, y) > 0$.

3. Give examples of joint random variables X, Y, Z such that
   - 3.1. $I(X : Y \mid Z) < I(X : Y)$,
   - 3.2. $I(X : Y \mid Z) > I(X : Y)$.

4. Let X and Y be random variables that take on values $x_1, x_2, \ldots, x_r$ and $y_1, y_2, \ldots, y_s$, respectively. Assume that these values are real numbers, so that we can add them. Let $Z = X + Y$.
   - 4.1. Show that if X and Y are independent, then $H(Z) \geq \max\{H(X), H(Y)\}$. Thus, the addition of *independent* random variables adds uncertainty.
   - 4.2. Give an example in which $H(Z) < \min\{H(X), H(Y)\}$.
   - 4.3. Under what conditions do we have $H(Z) = H(X) + H(Y)$? Make sure your condition is precise (i.e., necessary and sufficient) and give a thorough proof.

5. Let X, Y, Z be joint random variables. Prove the following inequalities and find conditions for equality.
   - 5.1. $H(XY|Z) \geq H(X|Z)$.
   - 5.2. $I(XY : Z) \geq I(X : Z)$.
   - 5.3. $H(XYZ) - H(XY) \leq H(XZ) - H(X)$.
   - 5.4. $I(X : Z \mid Y) \geq I(Z : Y|X) - I(Z : Y) + I(X : Z)$.

6. Suppose an instantaneous code over the binary alphabet $\{0, 1\}$ has codeword lengths $\ell_1, \ell_2, \ldots, \ell_m$. Suppose these lengths satisfy the Kraft inequality with some slackness, i.e., $\sum_{i=1}^{m} 2^{-\ell_i} < 1$. Prove that there exist arbitrarily long strings in $\{0, 1\}^*$ that cannot be decoded into sequences of codewords.

7. Let X and Y be random variables taking values in the finite sets $\mathcal{X}$ and $\mathcal{Y}$, respectively. Let $p(x) = \Pr[X = x]$ and

$$d(x) \;=\; \sum_{y \in \mathcal{Y}} \Big| \Pr[Y = y] - \Pr[Y = y \mid X = x] \Big| \,.$$

   Intuitively, $d(x)$ is a measure of how much the distribution of Y is affected when we are told that $X = x$. Based on this intuition, we can guess that the values $d(x)$ should be small if X does not reveal much information about Y, i.e., if $I(X : Y)$ is small. This is indeed the case. Prove that

$$\sum_{x \in \mathcal{X}} p(x) d(x) \;=\; O\left(\sqrt{I(X : Y)}\right) \,.$$

   Try to find the best possible constant that you can hide inside the big-O.