050/520/600.666 Information Extraction from Speech and Text

Homework # 6

Due April 7, 2005.

Entropy of discrete random variables

1. Starting with the definition

$$H(X_1, X_2, \ldots, X_n) = \sum_{x_1, x_2, \ldots, x_n \in \mathcal{X}^n} p(x_1, x_2, \ldots, x_n) \log \frac{1}{p(x_1, x_2, \ldots, x_n)} \,,$$

prove the chain rule of entropy:

$$H(X_1, X_2, \ldots, X_n) = H(X_1) + H(X_2|X_1) + \ldots + H(X_n|X_1, X_2, \ldots, X_{n-1}) \,.$$

2. Given two random variables X and Y, draw a Venn diagram to describe the relationships between $H(X,Y)$, $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$ and $I(X;Y)$.

3. Establish the relationship between the entropy $H(\mathbf{p})$ of a probability mass function $\mathbf{p}$ and the Kullback-Leibler distance $D(\mathbf{p}\|\mathbf{u})$ between $\mathbf{p}$ and the uniform distribution $\mathbf{u}$.

4. A fair coin is flipped until the first head occurs. Let X denote the number of flips required. Find the entropy of X in bits.

Conditional Mutual Information

Define the *conditional* mutual information between two random variables X and Y, given a third random variable Z, as

$$I(X;Y|Z) = \sum_{z \in \mathcal{Z}} p(z) I(X;Y|Z=z) = \sum_{z \in \mathcal{Z}} p(z) \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \,.$$

1. Manipulate the expression above to show that

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) \,.$$

2. From the expression above, derive the chain-rule of mutual information

$$I(X_1, X_2, \ldots, X_n; Y) = I(X_1;Y) + I(X_2;Y|X_1) + \ldots + I(X_n;Y|X_1, \ldots, X_{n-1}).$$

Coin Weighing Problems

Suppose one has n coins, among which there may or may not be one counterfeit coin. If there is a counterfeit coin, it may be either heavier or lighter than the other coins. The coins are to be weighed by a balance.

1. Find an upper bound on the number of coins n so that k weighings will find the counterfeit coin (if any) and declare it to be heavier or lighter.
2. What is the coin weighing strategy for $k = 3$ and $n = 12$ coins?