ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 10 Information Theory and Coding
Homework 5 October 18, 2011

PROBLEM 1. Show, for a Markov chain, that

$$H(X_0|X_n) \geq H(X_0|X_{n-1}), \quad n \geq 1.$$

Thus, initial state X_0 becomes more difficult to recover as time goes by.

PROBLEM 2. A discrete memoryless source emits a sequence of statistically independent binary digits with probabilities $p(1) = 0.005$ and $p(0) = 0.995$. The digits are taken 100 at a time, and a binary codeword is provided for every sequence of 100 digits containing three or fewer 1's.

(a) Assuming that all the codewords are the same length, find the minimum length required to provide codewords to all sequences with three or fewer ones.

(b) Calculate the probability of observing a source sequence for which no codeword has been assigned.

(c) Use Chebyshev's inequality to bound the probability of observing a source sequence for which no codewords has been assigned. Compare this bound to the actual probability computed in part (b).

PROBLEM 3. Let $X_1, X_2, \dots$ be i.i.d. random variables with distribution $p(x)$ taking values in a finite set $\mathcal{X}$. Thus, $p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i)$. We know that

$$-\frac{1}{n} \log p(X_1, \dots, X_n) \to H(X)$$

in probability. Let $q(x_1, \dots, x_n) = \prod_{i=1}^n q(x_i)$, where $q(x)$ is another probability distribution on $\mathcal{X}$.

(a) Evaluate

$$\lim_{n\to\infty} -\frac{1}{n} \log q(X_1, \dots, X_n).$$

(b) Now evaluate the limit of the log-likelihood-ratio

$$\frac{1}{n} \log \frac{q(X_1, \dots, X_n)}{p(X_1, \dots, X_n)}.$$

PROBLEM 4. Consider a valid, prefix-free dictionary of words from a source of alphabet size D.

(a) Show that the set of lengths $L_1, \dots, L_M$ of the dictionary words satisfy the Kraft inequality

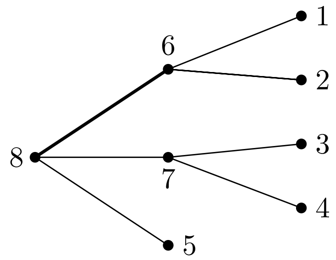
$$\sum_j D^{-L_j} \leq 1$$

with equality.

(b) Show that equality can happen only if the number of words $\mod (D-1) = 1$.

(c) Show that if the dictionary is valid, but not prefix-free, then the Kraft inequality is violated.

PROBLEM 5. Consider a tree with M leaves $n_1, \ldots, n_M$ with probabilities $P(n_1), \ldots, P(n_M)$. Each intermediate node n of the tree is then assigned a probability $P(n)$ which is equal to the sum of the probabilities of the leaves that descend from it. Label each branch of the tree with the label of the node that is on that end of the branch further away from the root. Let $d(n)$ be a "distance" associated with the branch labelled n. The distance to a leaf is the sum of the branch distances on the path to from root to leaf.



For example, in the tree shown above, nodes 1, 2, 3, 4, 5 are leaves, the probability of node 6 is given by $P(1) + P(2)$, the probability of node 7 by $P(3) + P(4)$, of node 8 (root) by $P(1) + P(2) + P(3) + P(4) + P(5) = 1$. The branch indicated by the heavy line would be labelled 6. The distance to leaf 2 is given by $d(6) + d(2)$.

(a) Show that the expected distance to a leaf is given by $\sum_n P(n)d(n)$ where the sum is over all nodes other than the root. Recall that we showed this in the class for $d(n) = 1$.

(b) Let $Q(n) = P(n)/P(n')$ where n' is the parent of n, and define the entropy of an intermediate node n' as

$$H_{n'} = \sum_{n:\ n \text{ is a child of } n'} -Q_n \log Q_n.$$

Show that the entropy of the leaves

$$H(\text{leaves}) = -\sum_{j=1}^{M} P(n_j) \log P(n_j)$$

is equal to $\sum_{n \in I} P(n)H_n$ where the sum is over all intermediate nodes including the root. Hint: use part (a) with $d(n) = -\log Q(n)$.

(c) Let X be a memoryless source with entropy H. Consider some valid prefix-free dictionary for this source and consider the tree where leaf nodes corresponds to dictionary words. Show that $H_n = H$ for each intermediate node in the tree, and show that

$$H(\text{leaves}) = E[L]H$$

where $E[L]$ is the expected word length of the dictionary. Note that we proved this result in class by a different technique.