6.450 Principles of Digital Communications
MIT, Fall 2004

Wednesday October 13, 2004
9-11am

Quiz 1

- This is a closed-book quiz, except that three 8.5″ × 11″ sheets of notes are allowed.
- There are three problems on the quiz of approximately equal value.
- The problems are not necessarily in order of difficulty.
- A correct answer does not guarantee full credit and a wrong answer does not guarantee loss of credit. You should concisely indicate your reasoning and show all relevant work. The grade on each problem is based on our judgment of your level of understanding as reflected by what you have written.
- If we can't read it, we can't grade it.
- If you don't understand a problem, please ask.
- Calculators are allowed, but probably won't be useful

Problem Q1.1 (35 points)

Consider a quaternary-valued source governed by a four-state Markov model that is operating in its steady state.

As shown in Figure 1, when in a state i, a transition to state $[(i+1) \mod 4]$ occurs with probabity p, where p is a parameter, and such a transition causes a source symbol $[(i+1) \mod 4]$ to be produced; when no transition occurs, the source symbol i is produced.

(a) Suppose the memory in the source is ignored and a simple Huffman code is used to code each symbol. What average rate R_1 in bits per source symbol is the source compressed to?

(b) Determine the average rate R in bits per source symbol achievable by the best code for this source if the memory is taken into account. Express your answer in terms of the binary entropy function $H_{\mathrm{B}}(\alpha) = -\alpha \log_2(\alpha) - (1-\alpha)\log_2(1-\alpha)$. Is there a value of p for which your answer is close to that of part (a)? Explain.

(c) Suppose the source is partitioned into consecutive length-2 blocks, inter-block statistical dependencies are ignored, and a 2-to-variable length code of mininum expected length is used, which exploits the intra-block dependencies. The achievable average rate in bits per source symbol in this case is guaranteed to be at least as low as $R_2 = aH_{\mathrm{B}}(b) + c$, where $H_{\mathrm{B}}(\cdot)$ is again the binary entropy function. Determine a, b, and c, expressing your answers in terms of the parameter p.

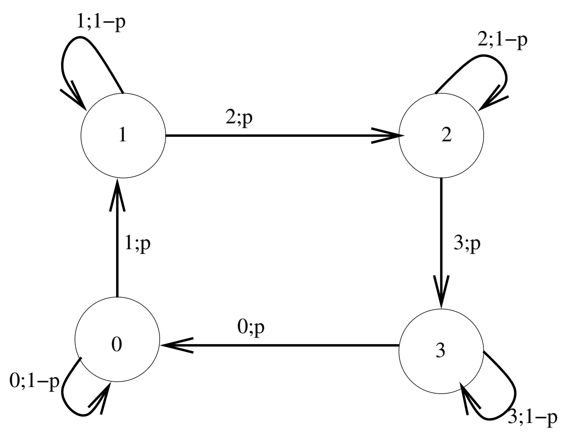

Figure 1: Markov Model for Problem Q1.1

Problem Q1.2 (35 points)

Consider encoding the set of positive integers and assume that a probability distribution on these integers satisfies $p_i > p_j$ for $i < j$. The table below gives the binary representation, the unary code and the unary-binary code for the first few positive integers. (Recall that the unary-binary code is used in LZ77. This code works by viewing an integer j as a pair of integers (n, m) where $n = \lfloor \log j \rfloor + 1$ and $m = j - 2^{n-1}$. Then j is mapped into the unary code for n concatenated with the $n-1$ bit binary representation for m.)

j	binary rep.	unary code	unary-binary code
1	1	1	1
2	10	01	010
3	11	001	011
4	100	0001	00100
5	101	00001	00101
6	110	000001	00110

(a) Does the binary representation constitute a prefix-free code? Explain. How about the unary and unary-binary codes?

(b) Show that if an optimum (in the sense of minimum expected length) prefix-free code is chosen for any given pmf (subject to the condition $p_i > p_j$ for $i < j$), the code word lengths satisfy $l_i \leq l_j$ for all $i < j$. Use this to show that for all $j \geq 1$

$$l_j \geq \lfloor \log j \rfloor + 1$$

(c) The asymptotic efficiency of a prefix-free code for the positive integers is defined to be $\lim_{j\to\infty} \frac{\log j}{l_j}$. What is the asymptotic efficiency of the unary-binary code?

(d) Explain how to construct a prefix-free code for the positive integers where the asymptotic efficiency is 1. Hint: Replace the unary code for the integers $n = \lfloor \log j \rfloor + 1$ in the unary-binary code with a code whose length grows more slowly with increasing n.

Problem Q1.3 (True or False) 30 points

For each of the following, state whether the statement is true or false and briefly indicate your reasoning. No credit will be given without a reason.

(a) Suppose X and Y are binary valued random variables with probability mass function given by $p_X(0) = 0.2$, $p_X(1) = 0.8$ $p_Y(0) = 0.4$ and $p_Y(1) = 0.6$. The joint probability mass function that maximizes the joint entropy $H(X, Y)$ is given by

$p_{X,Y}(\cdot,\cdot)$	X =0	X=1
Y=0	0.08	0.32
Y=1	0.12	0.48

(b) For a discrete memoriless source X with alphabet $\mathcal{X} = \{1, 2, \ldots, M\}$, let $L_{\min,1}$, $L_{\min,2}$, and $L_{\min,3}$ be the normalized average length in bits per source symbol for a Huffman code over $\mathcal{X}$, $\mathcal{X}^2$ and $\mathcal{X}^3$ respectively. Then there exists a specific probability mass function for source X for which $L_{\min,3} > \frac{2}{3}L_{\min,2} + \frac{1}{3}L_{\min,1}$.

(c) Assume that a continuous valued rv Z has a probability density that is 0 except over the interval $[-A, +A]$. Then the differential entropy $h(Z)$ is upper bounded by $1 + \log_2 A$. Also $h(Z) = 1 + \log_2 A$ if and only if Z is uniformly distributed between $-A$ and $+A$.