ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
Communication Systems Department

Handout 19 Information Theory and Coding
Midterm December 18, 2001

Three problems with a total of 115 points (100+bonus).
Duration 2 hours and 45 minutes.
4 sheets of notes allowed.

PROBLEM 1. (35 points.) Fano's inequality gives a lower bound on the probability of error in terms of the conditional entropy. In this problem we will derive an *upper* bound on the error probability of a maximum-aposteriori-probability (MAP) decoder in terms of the conditional entropy.

Let U and V be random variables with joint distribution p_{UV}. We will think of U as being the message, and V the received symbol. Let $\hat{U}$ be the MAP estimate of U from V.

Given that $V = v$ is received, the MAP decoder will decide on a $\hat{u}$ for which

$$\Pr(U = \hat{u}|V = v) \geq \Pr(U = u|V = v) \quad \text{for all } u.$$

(a) (5 points) Show that when $V = v$, the MAP decoder makes an error with probability

$$\Pr(U \neq \hat{U}|V = v) = 1 - \max_u p_{U|V}(u|v).$$

(b) (5 points) Show that

$$\Pr(\hat{U} \neq U) = \sum_v p_V(v)[1 - \max_u p_{U|V}(u|v)].$$

(c) (15 points) Show that for any random variable W,

$$\begin{aligned} H(W) &\geq \sum_w p_W(w)[1 - p_W(w)](\log e) \\ &\geq [1 - \max_w p_W(w)](\log e). \end{aligned}$$

[Hint: $\ln z \leq z - 1$.]

(d) (5 points) Show that

$$H(U|V = v) \geq [1 - \max_u p_{U|V}(u|v)](\log e).$$

(e) (5 points) Show that

$$\Pr(\hat{U} \neq U)(\log e) \leq H(U|V).$$

Problem 2. (35 points) Consider a binary, stationary, Markov source described by

$$\Pr(X_{k+1} = 0 \mid X_k = 0) = \Pr(X_{k+1} = 1 \mid X_k = 1) = \alpha$$

where $0 < \alpha < 1$.

(a) (5 points) Find the entropy rate of this source.

Given a sequence $X_1, X_2, \dots$, we can think of it as an alternating series of repetitions. For example if

$$X_1, X_2, \dots = 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, \dots$$

it can be thought of as 0 repeated 3 times, 1 repeated 2 times, 0 repeated 1 time, 1 repeated 5 times, etc. Let $R_1, R_2, \dots$ be the lengths of these repetitions. In the example above, these are 3,2,1,5,....

(b) (5 points) Argue that $R_1, R_2, \dots$, form an i.i.d. sequence.

(c) (7 points) Find the probability distribution of R_k.

(d) (10 points) Find the expectation $E[R_1]$, and the entropy $H(R_1)$. You may find the formula $\sum_{k=0}^{\infty} k\alpha^k = \alpha/(1-\alpha)^2$ useful.

(e) (8 points) The sequence $X_1, X_2, \dots$ can be described by first describing X_1 using 1 bit and then describing $R_1, R_2, \dots$. Suppose the sequence $R_1, R_2, \dots$ is efficiently encoded into $\mathcal{H}(R)$ bits per symbol.

How many bits per symbol does this method use in encoding the sequence $X_1, X_2, \dots$? How does this compare to $\mathcal{H}(X)$ found in (a)?

Problem 3. (45 points) Given a *stationary* source $X_1, X_2, \ldots,$ define

$$A_n(X) = \frac{1}{n} H(X_{n+1}, \ldots, X_{2n} | X_1, \ldots, X_n).$$

(a) (15 points) Show that $A_n(X) \geq A_{n+1}(X)$.

(b) (5 points) Find upper and lower bounds on $A_n(X)$ of the form

$$H(X_k | X_1, \ldots, X_{k-1}) \leq A_n(X) \leq H(X_m | X_1, \ldots, X_{m-1})$$

where k and m are appropriately chosen.

(c) (5 points) Show that

$$\lim_{n \to \infty} A_n(X) = \mathcal{H}(X)$$

where $\mathcal{H}(X)$ is the entropy rate of the source.

(d) (15 points) Given n, consider a source coding scheme that operates as follows:

1. The source output $X_1, X_2, \ldots,$ is parsed into blocks of length n. The first block is $Y_1 = (X_1, \ldots, X_n)$, the second $Y_2 = (X_{n+1}, \ldots, X_{2n})$, the third, $Y_3 = (X_{2n+1}, \ldots, X_{3n})$, etc. Let $\mathcal{Y}$ denote the set of values of Y_k.
2. The first block Y_1 is coded in some uniquely decodable way, e.g., by using $\lceil \log |\mathcal{Y}| \rceil$ bits.
3. For $k \geq 1$, the $k+1$st block Y_{k+1} is encoded into a codeword which may depend on Y_k. In other words, for each $v \in \mathcal{Y}$ we have a uniquely decodable code $C_v : \mathcal{Y} \to \{0,1\}^*$, and to encode Y_{k+1} we use the code C_{Y_k}.

Note that the long-term performance of this scheme will be determined by the performance of step 3.

Let L_n be the average number of bits per source letter emitted by step 3. Find a lower bound to L_n that applies to the scheme described above. [Hint: consider the quantity A_n.]

(e) (5 points) Find an upper bound on the minimum possible L_n.