ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 31 Information Theory and Coding
Final Exam Jan. 16, 2014

3 problems, 110 points
180 minutes
4 sheets of notes allowed

Good Luck!

PLEASE WRITE YOUR NAME ON EACH SHEET OF YOUR ANSWERS

PLEASE USE A SEPARATE SHEET TO ANSWER EACH QUESTION

Problem 1. (30 points) For a source with alphabet $\mathcal{U}$, two engineers have designed two dictionary based compression schemes. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be the dictionaries the engineers have designed. Both dictionaries are valid and prefix free. Suppose the dictionaries are of the same size M.

Based on the work of these engineers, we propose the following compression method: given the source sequence $U_1U_2U_3\dots$

(i) let $W_1 = U_1 \dots U_{L_1}$ be the first word engineer 1's dictionary produces, let $W_2 = U_1 \dots U_{L_2}$ be the first word engineer 2's design would have produced.

(ii) choose $W = U_1 \dots U_L$ to be either W_1 or W_2 (using some decision rule). [E.g., W is the shorter of the two; W is the longer of the two; W is chosen by a coin flip; ...]

(iii) describe W using a fixed number (say b) of bits.

(iv) repeat the above, by continuing to parse $U_{L+1}U_{L+2}\dots$.

(a) (10 pts) We need to ensure that b is large enough so that that description in (iii) uniquely specifies W. Give upper and lower bounds on b (in terms of M and $|\mathcal{U}|$). [Hint: for the lower bound consider the case where $\mathcal{U} = \{x, y, z\}$, $M = 9$, $\mathcal{D}_1 = \{x, y, zx, zy, zzx, zzy, zzzx, zzzy, zzzz\}$, $\mathcal{D}_2 = \{z, y, xz, xy, xxz, xxy, xxxz, xxxy, xxxx\}$ and that in (ii) we use the "shorter" rule. Generalize to arbitrary $\mathcal{U}$ and M.]

(b) (5 pts) Our design team suggests to use in (iii) a value of b that allows unique specification of W regardless of the rule is used in (ii). Give upper and lower bounds to this value of b (in terms of M).

Assume that this value of b is used for the rest of the problem.

(c) (5 pts) In our design team a conflict arises on how to perform the choice in (ii) so that we use the least number of bits per source letter: some in the team say that W should be the shorter of W_1 and W_2, others say it should be the longer. Who is right?

Suppose the dictionary $\mathcal{D}_1$ is constructed using the Tunstall method assuming that source had distribution $p_1(u)$. (similarly $\mathcal{D}_2$ is constructed by the Tunstall method assuming that the source had distribution $p_2(u)$.) Denote by H_1 and H_2 the two entropies of the two distributions. Keep in mind that the Tunstall procedure is asymptotically optimal, i.e., if p_1 is the true distribution

$$\frac{\log M}{E_1[\text{length}(W_1)]} \le H_1 + \epsilon_1$$

where $E_1[\cdot]$ denotes expectation under the assumption that p_1 is the true distribution, and $\epsilon_1 > 0$ can be made arbitrarily small by taking M large enough. (Similarly for p_2).

(d) (10 pts) Show that the proposed scheme described in (i)–(iv), with the correct design choices in (b) and (c) will ensure that for both $k = 1$ and $k = 2$

$$\frac{b}{E_k[\text{length}(W)]} \le H_k + \epsilon$$

where ϵ can be made arbitrarily small by taking M large.

PROBLEM 2. (50 points) An *binary encoder for a pair of messages* of blocklength n is a device which accepts a pair (m_1, m_2) of messages with $m_1 \in \{1, \dots, M_1\}$ and $m_2 \in \{1, \dots, M_2\}$ and produces a binary sequence $x^n(m_1, m_2)$ of length n. For such a device we define a pair of rates (R_1, R_2) with

$$R_1 = \frac{1}{n} \log M_1, \quad R_2 = \frac{1}{n} \log M_2.$$

Let W_1 and W_2 be independent messages, uniformly distributed on $\{1, \dots, M_1\}$ and $\{1, \dots, M_2\}$. Suppose the sequence $X^n = x^n(W_1, W_2)$ produced by the encoder is transmitted (i) over a memoryless binary erasure channel with erasure probability p, and also (ii) over a memoryless binary erasure channel with erasure probability q, with $q > p$. Denote by Y^n and Z^n the outputs of these erasure channels (Y^n is the output of the BEC(p) and Z^n is the output of the BEC(q)).

(a) (5 pts) Show that $I(W_1W_2; Z^n) \le n(1-q)$.

(b) (5 pts) Show that $I(W_1; Z^nW_2) \le n(1-q)$.

(c) (5 pts) Show that $I(W_1; Z^n) = I(W_1W_2; Z^n) - I(W_2; Z^nW_1)$.

Suppose now that the encoder is constructed randomly, i.e., the collection $x_1(1,1), \dots, x_n(M_1, M_2)$ of nM_1M_2 binary symbols that specify the encoder are chosen i.i.d. with equal probability of being '0' or '1'.

(d) (5 pts) Given an encoder, let Δ_n be the probability that Y^n does not uniquely determine (W_1, W_2). Explain why (or why not) the condition $R_1 + R_2 < 1 - p$ guarantees that $\lim_{n\to\infty} E[\Delta_n] = 0$. (The expectation is over the randomness in the choice of encoder.)

(e) (10 pts) Given an encoder let Θ_n be the probability that (Z_n, W_1) does not uniquely determine W_2. Explain why (or why not) the condition $R_2 < 1 - q$ guarantees that $\lim_{n\to\infty} E[\Theta_n] = 0$.

(f) (5 pts) Show that for any $R_1 < q - p$, $R_2 < 1 - q$ and $\epsilon > 0$ there exists a sufficiently large n, and an encoder for which

$$\Delta_n < \epsilon, \quad \Theta_n < \epsilon.$$

(g) (10 pts) Show that for the encoder in (f), $H(W_2|Z^nW_1) \le nR_2\epsilon + h_2(\epsilon)$, and

$$I(W_1; Z^n) \le n(1 - q - R_2) + nR_2\epsilon + h_2(\epsilon).$$

($h_2(\cdot)$ is the binary entropy function.)

(h) (5 pts) Show that for any $R_1 \le q - p$ and $\epsilon > 0$ there exists a sufficiently large n and an encoder for which

$$\Delta_n < \epsilon, \quad \frac{1}{n} I(W_1; Z^n) < \epsilon.$$

PROBLEM 3. (30 points) Suppose we are told that for any n and M, for any binary code with blocklength n with M codewords, the minimum distance $d_{\min}$ satisfies $d_{\min} \le d_0(M, n)$ where d_0 is a specified upper bound on minimum distance.

(a) (10 pts) Show that any such upper bound d_0 can be improved to the following upper bound: for any n, M, for any binary code with blocklength n with M codewords

$$d_{\min} \le d_1(M, n)$$

where $d_1(M, n) = \min_{k:\, 0 \le k \le n} d_0(\lceil M/2^k \rceil, n - k)$.

(b) (5 pts) Consider the trivial bound

$$d_0(M, n) = \begin{cases} n & M \ge 2 \\ \infty & M \le 1. \end{cases}$$

What is the bound d_1 constructed via (a) for this d_0?

(c) (5 pts) Suppose we are given a binary code with M codewords of blocklength n. Fix $1 \le i \le n$ and let $a_1, \ldots, a_M$ be the ith bits of the M codewords. Suppose M_1 of the a_m's are '1' and M_0 of them are '0'. Show that

$$\begin{aligned} \sum_{m=1}^{M} \sum_{\substack{m'=1 \\ m' \ne m}}^{M} d_H(a_m, a_{m'}) &= 2M_0M_1 \\ &\le M^2/2. \end{aligned}$$

(d) (5 pts) Show that for any binary code with $M \ge 2$ codewords $\mathbf{x}_1, \ldots, \mathbf{x}_M$ of blocklength n

$$M(M-1)d_{\min} \le \sum_{m=1}^{M} \sum_{\substack{m'=1 \\ m' \ne m}}^{M} d_H(\mathbf{x}_m, \mathbf{x}_{m'}) \le nM^2/2;$$

consequently $d_{\min} \le \lfloor \frac{1}{2} n \frac{M}{M-1} \rfloor$.

(e) (5 pts) Compare the bound above with its improvement via (a) for $n = 15$, $M = 1024$.