CSE522/Math581EB Coding Theory **Fall 2001**

Exercise 1: October 9

The exercises are for bonus points. Try to solve the ones which you find interesting.

1.1 The erasure channel capacity

Consider the binary erasure channels, where bits are not flipped, but only erased, so $\Sigma_{in}\{0,1\}$, $\Sigma_{out} = \{0,1,?\}$ and $P_{i,i} = 1-p$, $P_{i,?} = p$ for $i = 0,1$. Prove that the channel has capacity $1-p$, i.e. prove that

- When $R < 1-p$ there exist a pair of function $E : \{0,1\}^k \to \Sigma_{in}^{k/R}$ and $D : \Sigma_{out}^{k/R} \to \{0,1\}^k$, with the following property. If we pick a message uniformly at random from $\{0,1\}^k$, encode it with E, send the result through the noisy channel and decode the result, then we recover the original message with probability $1-o(1)$,
$$\mathbf{P}[D(\mathcal{C}(E(x))) = x] = 1 - o(1).$$
- For all $R > 1-p$ and all functions $E : \{0,1\}^k \to \Sigma_{in}^{k/R}$ and $D : \Sigma_{out}^{k/R} \to \{0,1\}^k$, it holds that if we pick a message at random from $\{0,1\}^k$, encode it with E, send the result through the noisy channel and decode the result, then we recover the original message with probability $o(1)$,
$$\mathbf{P}[D(\mathcal{C}(E(x))) = x] = 1 - o(1).$$

This should be quite easy (but not very short), if you understand the proof for the binary symmetric channel.

1.2 Entropy and the noiseless coding theorem

. The entropy of a distribution $p = (p_1, \ldots, p_n)$ is defined by

$$H(p) = -\sum_{i=1}^{n} p_i \log_2 p_i.$$

In the exercise below we show how $H(p)$ is related to the average length of a good encoding of p.

- Let $(\ell(i))_{i=1}^n$ be some integer numbers. Then there exists a set of binary strings $(s_i)_{i=1}^n$ s.t. for all i, the length of s_i, denoted $|s(i)|$ is $\ell(i)$, and for all $i \neq j$, s_i is not a prefix of s_j **iff** $\sum_{i=1}^{\ell} 2^{-\ell(i)} \leq 1$. (Not hard: Try induction, or look at a book). The reason for looking at such sets of strings is that given a sequence of such strings it is possible to find the beginning and ends of the strings.
- Let $p = (p_1, \ldots, p_n)$ be a distribution. Then there exist a set of binary strings $s_1, \ldots, s_n$ such that for all i and j, s_i is not a prefix of s_j, and $\sum_{i=1}^n p_i|s_i| \in [H(p), H(p)+1]$. In other words, it is possible to encode the distribution with strings s.t. the average length of a string is in $[H(p), H(p)+1]$. (Hint : consider first the case where for all i, $p_i = 2^{-\ell(i)}$).
- Prove that there exists no such encoding where $\sum_{i=1}^n p_i|s_i| < H(p)$. (Hint: You'll have to use some properties of the entropy function, the buzz-word here is convexity, so keep in mind that $\log x \leq x-1$).

1.3 Hamming codes

- Construct Hamming codes over the Field $\mathcal{F}_q$ where $q > 2$.
- Suppose that an $(n, k, 2t + 1)_2$ code exists. Show that an $(n, k, 2t + 2)_2$ code exists.
- Is the last statement true for values of q different from 2? (I didn't think about it yet; it may be hard)

1.4 What's the color of your hat?

15 People are standing in a circle - each one is assigned a hat which is red/blue with probability 1/2 independently for all people. Each person sees the colors of all the other hats but not his own. After the hats are assigned there is no communication between the people. After the hats are assigned each of the people should declare the color of his hat, or say pass (the declarations are done simultaneously, so this is not a way to communicate). The group wins if at least one of the people declares a color, and all the people who declared a color, declared the color of their hat; otherwise, they lose. Find the best strategy the group has in order to win. (**Hints:** try the case of 3 people first. What is this problem doing in a coding course? **Warning:** This is not easy and quite addictive).