

## Math 345 - Problem Set 4 - due Wednesday, April 18

These will be graded on both correctness and clarity, so write careful, organized solutions (a good idea to keep in mind is that another person from the class should be able to read and understand your solution). All of your answers must be justified. **Honor Code:** For this and all problem sets, you are encouraged to work on solving these problems in groups. However, you **MUST** write up your solutions individually; in particular, you may not look at someone else's write-up. In addition, you must indicate who you worked with.

1. Returning to problem 4b from the midterm (sending the letters  $a, b, c, d$  using a Morse code where a dot/0 takes 1 time unit and a dash/1 takes 3 time units), for the particular probability distribution on  $\{a, b, c, d\}$  given in the problem, the codewords  $\{000, 001, 01, 1\}$  were optimal among all prefix codes (for some particular assignment of  $a, b, c, d$  to them).
  - (a) List out the 5 different sets of possible codewords that make complete prefix codes (hint: 3 of the 5 are already listed in this problem).
  - (b) Show that for some different probability distribution, the codewords  $\{0, 10, 110, 111\}$  are optimal.
  - (c) Is there a probability distribution such that  $\{0, 100, 101, 11\}$  are optimal?

(Note: be more careful in your explanations than I let you be on the midterm).

2. Define the code  $R_0 = \{0, 1\} \subset \mathbb{Z}_2^1$ , and recursively define  $R_m \subset \mathbb{Z}_2^{2^m}$ , for  $m \geq 1$ , by

$$R_m = \{(u, u) : u \in R_{m-1}\} \cup \{(u, u + \mathbf{1}) : u \in R_{m-1}\},$$

where  $\mathbf{1}$  is the string of all 1's and “ $(u, u)$ ” means concatenate  $u$  with itself. For example

$$R_1 = \{00, 11\} \cup \{01, 10\} = \{00, 11, 01, 10\}.$$

- (a) List out the elements of  $R_3$ .
- (b) Prove that, for all  $m$ ,
- $|R_m| = 2^{m+1}$ ,
  - $R_m$  is a linear code (hint: use the definition  $u, v \in C \Rightarrow u + v \in C$ ), and
  - every codeword of  $R_m$  except  $00 \cdots 0$  and  $11 \cdots 1$  has weight exactly  $2^{m-1}$ .
- (c) Conclude that  $R_m$  has minimum distance  $2^{m-1}$  (and so it can correct  $2^{m-2} - 1$  errors).

These are the  $R(1, m)$  *Reed-Muller codes* (a more general recursive definition can define all  $R(n, m)$ ). They used to be very popular in situations where there was the necessity of correcting a lot of errors. For example  $R_5$  was used to communicate with the Mariner spacecraft in the 60's and 70's. They have since been replaced with more efficient — but more complicated — codes like *Reed-Solomon codes* (which we may talk a little about later).

3. Recall that we create the Hamming code  $H_m \subset \mathbb{Z}_2^{2^m-1}$  by first indexing the bits of each string  $\mathbf{b}$  in  $\mathbb{Z}_2^{2^m-1}$  by  $b_A$ , where  $A$  is a nonempty subset of  $\{1, 2, \dots, m\}$  and then defining  $H_m$  to be the strings such that, for all  $i \in \{1, 2, \dots, m\}$ ,

$$\sum_{A \subset \{1, 2, \dots, m\}: i \in A} b_A = 0$$

(in  $\mathbb{Z}_2$ ). For example  $H_3 \subset \mathbb{Z}_2^7$  is the set of strings

$$\mathbf{b} = b_{\{1\}}b_{\{2\}}b_{\{3\}}b_{\{1,2\}}b_{\{1,3\}}b_{\{2,3\}}b_{\{1,2,3\}}$$

such that

$$\begin{aligned} b_{\{1\}} + b_{\{1,2\}} + b_{\{1,3\}} + b_{\{1,2,3\}} &= 0, \\ b_{\{2\}} + b_{\{1,2\}} + b_{\{2,3\}} + b_{\{1,2,3\}} &= 0, \text{ and} \\ b_{\{3\}} + b_{\{1,3\}} + b_{\{2,3\}} + b_{\{1,2,3\}} &= 0. \end{aligned}$$

- (a) Suppose that  $\mathbf{b}$  is a codeword in  $H_m$  of weight 3 (we proved in class that 3 is the minimum weight of  $H_m$ ). Suppose further that  $b_A = 1$  and  $b_B = 1$  for some  $A \neq B$ . Show that the third nonzero bit is uniquely determined (you should be able to give a formula for which bit it is in terms of  $A$  and  $B$ ).
- (b) Conclude that there are exactly

$$\frac{1}{3} \binom{2^m - 1}{2} = \frac{(2^m - 1)(2^m - 2)}{6}$$

codewords in  $H_m$  of weight 3.

4. (a) Suppose we want a code such that
- if 1 error is made, it can correct it, but also
  - if 2 errors are made, it can detect that at least 2 errors have been made (and so not accidentally “correct” it to the wrong codeword).

What minimum distance do we need to guarantee that our code has this property?

- (b) Suppose we take the Hamming code  $H_m \subset \mathbb{Z}_2^{2^m-1}$  and add a parity check bit  $b_\Omega$ , so that the sum of all of the bits is 0. Call this new code  $H'_m \subset \mathbb{Z}_2^{2^m}$ . For example  $H'_3 \subset \mathbb{Z}_2^8$  is the set of strings

$$\mathbf{b}' = b_{\{1\}}b_{\{2\}}b_{\{3\}}b_{\{1,2\}}b_{\{1,3\}}b_{\{2,3\}}b_{\{1,2,3\}}b_\Omega$$

such that

$$b_{\{1\}} + b_{\{1,2\}} + b_{\{1,3\}} + b_{\{1,2,3\}} = 0,$$

$$b_{\{2\}} + b_{\{1,2\}} + b_{\{2,3\}} + b_{\{1,2,3\}} = 0,$$

$$b_{\{3\}} + b_{\{1,3\}} + b_{\{2,3\}} + b_{\{1,2,3\}} = 0, \text{ and}$$

$$b_{\{1\}} + b_{\{2\}} + b_{\{3\}} + b_{\{1,2\}} + b_{\{1,3\}} + b_{\{2,3\}} + b_{\{1,2,3\}} + b_\Omega = 0.$$

Prove that  $H'_m$  has the minimum distance required in part (a).