MAT 312/AMS 351
Applied Algebra
Midterm 1 – Solutions

March 2, 2009

Note: There were 2 similar exams; this is one of them.

1. Consider the binary code with the following 6 code words:

$$\begin{array}{cc} U & 1001000 \\ V & 0111100 \\ W & 1001101 \\ X & 0101011 \\ Y & 1100100 \\ Z & 1110001 \end{array}$$

(a) (10 points) Is this a group code? Explain your answer.

SOLUTION: No. A group code must contain the zero word. Alternatively, $U+V = 1110100$ is not in the code, etc.

(b) (10 points)
Fill in the Hamming distances between the code words:

SOLUTION:

$H(\ ,\ )$	U	V	W	X	Y	Z
U	0	5	2	4	4	4
V		0	4	4	3	4
W			0	4	3	4
X				0	5	4
Y					0	3
Z						0

(c) (10 points) Can this code be used for single error detection? Explain your answer.

SOLUTION: Yes, because the minimum Hamming distance between different words is 2.

(d) (10 points) Show how one of the code words can be changed to make this code suitable for single-error correction.

SOLUTION: Changing U to 1001010 for example will change the table to

$H(\ ,\ )$	U	V	W	X	Y	Z
U	0	6	3	3	5	5
V		0	4	4	3	4
W			0	4	3	4
X				0	5	4
Y					0	3
Z						0

which now has mininum Hamming distace 3, and so is suitable for single-error correction.

(e*) (10 points) Your modified code uses 7 bits to encode the 6 symbols U, V, W, X, Y, Z. Could a shorter code (fewer bits) also allow single error correction?

SOLUTION: The canonical check-bit matrix

$$K = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

defines by its null-space a group code suitable for single error correction (since no zero column and no two columns equal). This will be a six-bit code with 3 check bits and 3 data bits, so it will allow up to 8 different symbols to be encoded.

2. Consider the matrix

$$M = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(a) (10 points) Explain why the null space of this matrix is a group code and suitable for single error correction.

SOLUTION: The null space of M is the set of 7-bit words w such that $Mw^t = 0$. If w and z are in the null space, then $M(w+z)^t = M(w^t + z^t) = Mw^t + Mz^t = 0$, i.e. the sum of 2 code words is a code word; so this is a group code. It is suitable for single-error correction because no column has all zeroes and no 2 columns are equal.

(b) (10 points) How many words are there in this code?

SOLUTION: There are 4 check bits and 3 data bits, so $2^3 = 8$ code words.

(c*) (10 points) Could this code be used for double-error detection AND single-error detection? Explain your answer.

SOLUTION: For double error detection and single-error *correction* (what the text should have been) a code most satisfy $d \geq 4$. This means that no word with $d \leq 3$ can be in the null space. We know how to guarantee no Hamming-distance 1 pairs (no zero column) and no Hamming-distance 2 pairs (no 2 columns equal). If w and v have Hamming distance 3 and are in the null-space, then their sum (a word with exactly 3 ones) is also in the null-space. M times (the transpose of) a word with exactly 3 ones is the sum of 3 columns of M. So to make this impossible we need to show that no three columns of M sum to zero. This can be checked case-by-case, or one can argue as follows: Any 3 columns must have 2 from the first 3 OR 2 from the last 4. Adding two of the first 3 columns gives a 4-vector with two ones. It can't be canceled by the third of the first three (has 3 ones) or by one of the last 4 (have one 1). Adding two of the last four also gives a 4-vector with 2 ones, which can't be canceled by one of the first three or by one of the other last four. So there is no way to add three columns to get zero.

3. (a) (10 points) Calculate the entropy E associated with the set of probabilities: $\{1/4, 1/4, 1/8, 1/8, 1/8, 1/16, 1/16\}$. Show all your work.

SOLUTION: $E = -\sum_i p_i \log_2 p_i$. So $E = -2(1/4)\log_2(1/4) - 3(1/8)\log_2(1/8) - 2(1/16)\log_2(1/16)$
$= -2(1/4)(-2) - 3(1/8)(-3) - 2(1/16)(-4) = 1 + 9/8 + 1/2 = 21/8.$

(b) (10 points) Suppose A, B, C, D, E, F, G are the seven symbols in an alphabet, and that they occur with frequencies $\{1/4, 1/4, 1/8, 1/8, 1/8, 1/16, 1/16\}$ respectively. Give a binary code for this alphabet such that no code word appears as the beginning of another (a prefix code) and such that the average length of a word is exactly E.

SOLUTION: Following the algorithm, we split (AB) from (CDEFG), then A from B and (CD) from (EFG), then C from D and E from FG, and finally F from G. The code is then

A-00

B-01

C-100

D-101

E-110

F-1110

G-1111

with average length $2(1/4)(2) + 3(1/8)(3) + 2(1/16)(4) = 21/8 = E$

END OF EXAMINATION