Name: .... Student ID: ....

UNSW SCHOOL OF MATHEMATICS AND STATISTICS

MATH3411 INFORMATION CODES AND CIPHERS

2014 S2 **TEST 1** **VERSION A**

- Time Allowed: **45 minutes**

For multiple choice questions, **circle the correct answer**; each multiple choice question is worth **1 mark**. For written answer questions, **use extra paper**. Staple all papers together when finished.

1. You are given the following 7-bit ASCII codewords:

0	0110000	1	0110001	2	0110010	3	0110011
4	0110100	5	0110101	6	0110110	7	0110111

Define a 4-character 8-bit ASCII burst code by encoding characters in blocks of three together with a 4th character which is used as a check codeword. (This is similar to the 9-character 8-bit and 8-character 8-bit ASCII code studied in lectures.)

The message 00110011 10110010 00110111 10110100 is received but contains a single error.

What is the corrected and decoded message?

(a) 127 (b) 307 (c) 325 (d) 327 (e) None of these.

For the next 3 **questions**, let C be a binary linear code with check matrix

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Assume that the information bits correspond to columns 3, 4, 6, and 7.

2. The codeword encoding the message 1101 in code C is

(a) 0111101 (b) 1011101 (c) 1101101 (d) 1101011 (e) 1101101

3. A generator matrix G corresponding to the check matrix H for the code C has size

(a) 4×7 (b) 3×7 (c) 4×3 (d) 3×4 (e) None of these

4. The minimum distance $d(C)$ of the code C is

(a) 0 (b) 1 (c) 2 (d) 3 (e) 4

5. Consider a binary channel with bit-error probability p, where errors in different positions are independent. Suppose that a codeword $\mathbf{x}$ is sent from a binary repetition code with codewords of length 4. The probability that one or more errors occur and are detected but cannot be corrected is

(a) $4p^3(1-p)$ (b) $6p^2(1-p)^2$ (c) p^4 (d) $4p^3(1-p)+p^4$ (e) $6p^2(1-p)^2+p^4$

6. Let C be the binary linear code with basis $\{0101100, 1001010, 1011001\}$.
How many codewords are there in C?

(a) 3 (b) 4 (c) 8 (d) 16 (e) 128

7. The message $s_2s_4s_2s_1$ was encoded using a comma code of length 4.
The encoded message is

(a) 010001011 (b) 101110100 (c) 11011111100 (d) 0100010101 (e) 101110110

8. A binary UD-code has codewords lengths (not necessarily in order) 1, 2, 3, 5, ℓ.
What is the smallest value of ℓ for which the code exists?

(a) $\ell = 1$ (b) $\ell = 2$ (c) $\ell = 3$ (d) $\ell = 4$ (e) $\ell = 5$

9. Consider a compression code with codewords $\mathbf{c}_1 = 1$, $\mathbf{c}_2 = 01$, $\mathbf{c}_3 = 100$, $\mathbf{c}_4 = ?$, where $\mathbf{c}_4$ is to be chosen from the list of four possibilities below.
Which choice, if any, of $\mathbf{c}_4$ makes the resulting code uniquely decodable?

(a) $\mathbf{c}_4 = 0$ (b) $\mathbf{c}_4 = 011$ (c) $\mathbf{c}_4 = 000$ (d) $\mathbf{c}_4 = 001$ (e) None of these

10. Let $S = \{s_1, s_2\}$ be a source with probabilities $p_1 = \frac{1}{5}$ and $p_2 = \frac{4}{5}$.
The average length per original symbol of a **radix 3** Huffman code for the **second extension** S^2 of this source (constructed with the usual strategies) is

(a) $\dfrac{3}{5}$ (b) $\dfrac{39}{50}$ (c) $\dfrac{34}{25}$ (d) $\dfrac{34}{50}$ (e) $\dfrac{6}{5}$

11. [**5 marks**]

(a) Show that there is no uniquely decodable **ternary** (i.e. radix 3) code with codeword lengths $1, 2, 2, 2, 2, 2, 2, 3$, respectively.

(b) The symbol s_1 of the source $S = \{s_1, s_2\}$ occurs with probability $5/7$ and symbol s_2 occurs with probability $2/7$. Find a uniquely decodable binary code of minimal average length for S^2, assuming that successive symbols occur independently, and state the average length per original source symbol of the code.

Name: .... Student ID: ....

UNSW SCHOOL OF MATHEMATICS AND STATISTICS

MATH3411 INFORMATION CODES AND CIPHERS

2014 S2 **TEST 1** **VERSION B**

- Time Allowed: **45 minutes**

For multiple choice questions, **circle the correct answer**; each multiple choice question is worth **1 mark**. For written answer questions, **use extra paper**. Staple all papers together when finished.

1. There may be an error in the check digit in the ISBN number 0-19-061133-X. The correct check digit is

(a) 2 (b) 4 (c) 6 (d) 8 (e) None of these.

For the next 3 **questions**, let C be a binary linear code with check matrix

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Assume that the check bits correspond to columns 1, 2, 3, and 7.

2. The codeword encoding the message 101 in code C is

(a) 1011101 (b) 0011010 (c) 1101010 (d) 1110101 (e) 1111010

3. A generator matrix G corresponding to the check matrix H for the code C has size

(a) 3×7 (b) 4×7 (c) 4×3 (d) 3×4 (e) None of these

4. The minimum distance $d(C)$ of the code C is

(a) 0 (b) 1 (c) 2 (d) 3 (e) 4

5. Let C be a binary 1-error correcting code with k information bits, $m = 3$ check bits and 2^k codewords. The maximum possible value for k is

(a) 1 (b) 2 (c) 3 (d) 4 (e) 5

6. Let C be the code consisting of all vectors $\mathbf{x} = x_1x_2x_3x_4 \in \mathbb{Z}_5^4$ satisfying the check equations

$$\begin{array}{rcrcrcrcl} x_1 & + & x_2 & + & x_3 & & & \equiv & 0 \pmod 5 \\ x_1 & + & 3x_2 & & & + & 2x_4 & \equiv & 0 \pmod 5 \end{array}$$

Which, if any, of the following is a valid code word?

(a) 1122 (b) 2121 (c) 4341 (d) 3344 (e) None of these

7. The message $s_2s_1s_4s_3$ was encoded using a comma code of length 4.
The encoded message is

(a) 1011110110 (b) 1011111110 (c) 1101011110 (d) 0111110010 (e) 1001110110

8. A binary UD-code with minimal average codeword length has codeword lengths (not necessarily in order) 2, 2, 3, 3, 4, 4, ℓ. What is the value of ℓ?

(a) $\ell = 1$ (b) $\ell = 2$ (c) $\ell = 3$ (d) $\ell = 4$ (e) None of these

9. Consider a binary Huffman code for a source with 8 symbols $S = \{s_1, \ldots, s_8\}$, where the source symbols are given in non-increasing probability order. Suppose that the codeword for symbol s_7 is $\mathbf{c}_7 = 1100$. Then the codeword for symbol s_8 is

(a) 010 (b) 110 (c) 1101 (d) 1110 (e) 1111

10. Let $S = \{s_1, s_2, s_3, s_4, s_5\}$ be a source with probabilities $p_1 = \frac{2}{5}$, $p_2 = \frac{1}{5}$, $p_3 = \frac{1}{5}$, $p_4 = \frac{2}{15}$, $p_5 = \frac{1}{15}$. The average length of a **radix 3** Huffman code for this source (using the usual strategies) is

$(a)\ \dfrac{13}{15}$ $(b)\ \dfrac{14}{15}$ $(c)\ \dfrac{6}{5}$ $(d)\ \dfrac{22}{15}$ $(e)\ \dfrac{7}{5}$

11. **[5 marks]**

(a) Find an instantaneous **ternary** (i.e. radix 3) UD-code for the source

$$S = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8\}$$

with codeword lengths $1, 2, 2, 2, 2, 3, 3, 4$, respectively.

(b) The symbol s_1 of the source $S = \{s_1, s_2\}$ occurs with probability $3/5$ and s_2 occurs with probability $2/5$. Find a binary UD-code of minimal average length for S^2, assuming that successive symbols occur independently, and state the average length per original source symbol of the code.