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

UNSW School of Mathematics and Statistics

MATH3411 Information Codes and Ciphers

2016 S2 **TEST 2** **VERSION A**

- Time Allowed: **45 minutes**

For the multiple choice questions, **circle the correct answer**;
each multiple choice question is worth **1 mark**.
For the true/false and written answer questions, use extra paper.
Staple everything together at the end.

1. Using the LZ78 algorithm a message is encoded as $(0,a)(0,b)(2,a)(3,b)(4,a)$.
What is the last dictionary entry after decoding?

(a) *abab* (b) *abba* (c) *baab* (d) *aabb* (e) *baba*

2. A Markov source $S = \{s_1, s_2, s_3\}$ has transition matrix M. The Huffman code for the equilibrium distribution is $\text{Huff}_E = [1, 00, 01]$. (so $\mathbf{c}_1 = 1$, $\mathbf{c}_2 = 00$ and $\mathbf{c}_3 = 01$.)
Huffman codes for the columns of M are given by
$\text{Huff}_1 = [00, 1, 01]$, $\text{Huff}_2 = [0, 10, 11]$ and $\text{Huff}_3 = [11, 10, 0]$.
The string 001101100 decodes under the Markov Huffman encoding as
(a) $s_2s_1s_1s_3s_1s_1$ (b) $s_2s_3s_3s_1s_1$ (c) $s_3s_3s_1s_3s_2s_2s_2$ (d) $s_2s_2s_1s_2s_3$ (e) none of these

3. Consider a binary channel with source symbols $\{a_1, a_2\}$ and output symbols $\{b_1, b_2\}$ such that $P(a_1) = \frac{2}{5}$, $P(b_1 \mid a_1) = \frac{5}{8}$ and $P(b_2 \mid a_2) = \frac{5}{7}$. Recall the function

$$H(x) = -x\log_2 x - (1-x)\log_2(1-x)$$

and note that $H(x) = H(1-x)$. The noise entropy $H(B \mid A)$ can be written as

(a) $\frac{2}{5}H(\frac{3}{8})+\frac{3}{5}H(\frac{5}{7})$ (b) $\frac{2}{7}H(\frac{2}{5})+\frac{5}{8}H(\frac{3}{5})$ (c) $\frac{5}{8}H(\frac{2}{5})+\frac{2}{7}H(\frac{3}{5})$ (d) $\frac{3}{5}H(\frac{5}{8})+\frac{2}{5}H(\frac{2}{7})$ (e) $\frac{2}{7}H(\frac{3}{5})+\frac{5}{8}H(\frac{2}{5})$

4. Using Euler's Theorem or otherwise, calculate $5^{1155} \pmod{2016}$.
The answer is

(a) 1 (b) 5 (c) 25 (d) 125 (e) 625

5. For which of the following numbers a is $n = 121$ a strong pseudo-prime to base a?

(a) 2 (b) 3 (c) 5 (d) 7 (e) None of these

6. **[5 marks]** For each of the following, say whether the statement is true or false, giving a brief reason or showing your working. You will get $\frac{1}{2}$ **mark** for a correct true/false answer, and if your true/false answer is correct, then you will get $\frac{1}{2}$ **mark** for a good reason.

Begin each answer with the word "True" or "False".

i) If arithmetic coding with source symbols a, b and stop symbol $\bullet$ corresponding to the intervals $[0, 0.3)$, $[0.3, 0.7)$ and $[0.7, 1)$ is used, then the message $bb\bullet$ is encoded by 0.6.

ii) For a 3-symbol source $S = \{s_1, s_2, s_3\}$ with probabilities $p_1 = 4/9$, $p_2 = 2/9$, $p_3 = 1/3$ it is possible to find a **ternary** encoding of some extension S^n with average word length per original source symbol less than 1.

iii) When using Fermat factorisation to factor $n = 1333$ as a product $n = ab$ where $2 \leq a < b$, the sum $a + b$ equals 74.

iv) For a source $S = \{a, b\}$ with probabilities $P(a) = \frac{1}{3}$ and $P(b) = \frac{2}{3}$, the second longest codewords in the binary Shannon-Fano code for the fifth extension S^5 have length 7.

v) There are 100 primitive elements in the field GF(125).

7. **[5 marks]** Let $\mathbb{F} = \mathbb{Z}_3(\alpha)$ where α is a root of the polynomial $x^2 + 2x + 2 \in \mathbb{Z}_3[x]$.

(i) Express all non-zero elements of $\mathbb{F}$ as powers of α and as linear combinations over $\mathbb{Z}_3$ of 1 and α.

(ii) Solve the set of linear equations

$$\begin{pmatrix} \alpha^2 & \alpha^4 \\ \alpha & \alpha^6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha^2 \\ 1 \end{pmatrix}$$

in $\mathbb{F}$.

(iii) Find the minimal polynomial of α^5.
Show your working.

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

UNSW School of Mathematics and Statistics

MATH3411 Information Codes and Ciphers

2016 S2 **TEST 2** **VERSION B**

- Time Allowed: **45 minutes**

For the multiple choice questions, **circle the correct answer**;
each multiple choice question is worth **1 mark**.
For the true/false and written answer questions, use extra paper.
Staple everything together at the end.

1. Using the LZ78 algorithm a message is encoded as $(0,a)(0,b)(2,a)(2,b)(4,b)$.
What is the last dictionary entry after decoding?

(a) *aba* (b) *abb* (c) *bab* (d) *bba* (e) *bbb*

2. A Markov source $S = \{s_1, s_2, s_3\}$ has transition matrix M. The Huffman code for the equilibrium distribution is $\text{Huff}_E = [1, 00, 01]$. (so $\mathbf{c}_1 = 1$, $\mathbf{c}_2 = 00$ and $\mathbf{c}_3 = 01$.)
Huffman codes for the columns of M are given by
$\text{Huff}_1 = [00, 1, 01]$, $\text{Huff}_2 = [0, 10, 11]$ and $\text{Huff}_3 = [11, 10, 0]$.
Given the string of source symbols $s_1s_2s_1s_3s_1$, the Markov Huffman encoding is

(a) 0100110 (b) 111011011 (c) 1001011 (d) 1100111 (e) None of these

3. Consider a binary channel with source symbols $\{a_1, a_2\}$ and output symbols $\{b_1, b_2\}$ such that $P(a_1) = \frac{1}{4}$, $P(b_1 \mid a_1) = \frac{3}{8}$ and $P(b_2 \mid a_2) = \frac{2}{7}$. Recall the function

$$H(x) = -x \log_2 x - (1-x)\log_2(1-x)$$

and note that $H(x) = H(1-x)$. The noise entropy $H(B \mid A)$ can be written as

(a) $\frac{2}{7}H(\frac{1}{4})+\frac{3}{8}H(\frac{3}{4})$ (b) $\frac{3}{4}H(\frac{3}{8})+\frac{1}{4}H(\frac{2}{7})$ (c) $\frac{1}{4}H(\frac{3}{8})+\frac{3}{4}H(\frac{5}{7})$ (d) $\frac{3}{8}H(\frac{1}{4})+\frac{2}{7}H(\frac{3}{4})$ (e) $\frac{2}{7}H(\frac{3}{4})+\frac{3}{8}H(\frac{1}{4})$

4. Using Euler's Theorem or otherwise, calculate $2^{2016} \pmod{123}$. The answer is

(a) 4 (b) 9 (c) 25 (d) 36 (e) 100

5. For which of the following numbers a is $n = 25$ a strong pseudo-prime to base a?

(a) 2 (b) 3 (c) 5 (d) 7 (e) None of these

6. [**5 marks**] For each of the following, say whether the statement is true or false, giving a brief reason or showing your working. You will get $\frac{1}{2}$ **mark** for a correct true/false answer, and if your true/false answer is correct, then you will get $\frac{1}{2}$ **mark** for a good reason.

Begin each answer with the word "True" or "False".

i) If arithmetic coding with source symbols a, b and stop symbol $\bullet$ corresponding to the intervals $[0, 0.4)$, $[0.4, 0.8)$ and $[0.8, 1)$ is used, then the message $ba\bullet$ is encoded by 0.55.

ii) For a 3-symbol source $S = \{s_1, s_2, s_3\}$ with probabilities $p_1 = 1/2$, $p_2 = 1/6$, $p_3 = 1/3$, it is possible to find a **ternary** encoding of some extension S^n with average word length per original source symbol less than 0.9.

iii) When using Fermat factorisation to factor $n = 2257$ as a product $n = ab$ where $2 \leq a < b$, the linear combination $2a - b$ equals 13.

iv) A source $S = \{s_1, s_2\}$ has probabilities $P(s_1) = \frac{1}{4}$, $P(s_2) = \frac{3}{4}$. The second shortest codeword length in the binary Shannon-Fano code for the fourth extension S^4 is 3.

v) Given that 5 is a primitive element of $\mathbb{Z}_{17}$, then 6 is also a primitive element.

7. [**5 marks**] Let $\mathbb{F} = \mathbb{Z}_3(\alpha)$ where α is a root of the polynomial $x^2 + x + 2 \in \mathbb{Z}_3[x]$.

(i) Express all nonzero elements of $\mathbb{F}$ as powers of α and as linear combinations over $\mathbb{Z}_3$ of 1 and α.

(ii) Solve the set of linear equations

$$\begin{pmatrix} \alpha^3 & \alpha^4 \\ \alpha^4 & \alpha^6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha^2 \\ \alpha^5 \end{pmatrix}$$

in $\mathbb{F}$.

(iii) Find the minimal polynomial of α^2.
Show your working.