MATH3411 Information, Codes And Ciphers

2022 T3

Problems and Answers

MATH3411 Problems

Chapter 1: Introduction

1 **a.** Explain why, if n is not a prime, then n has a prime factor less than or equal to $\sqrt{n}$.

b. Explain why, for $18 \leq n \leq 400$, $n \neq 361$, that n is prime if and only if n does not have a proper factor 2 or 5 and $\gcd(n, 51051) = 1$.

c. Hence test $n = 323$ and $n = 373$ for primeness.

***2** Prove that if $2^n - 1$ is a prime, then n is prime and that if $2^n + 1$ is prime, then n is a power of 2.
This problem is a nice but hard challenge; feel free to skip!

3 **a.** What is the probability of picking a Queen from a standard pack of cards?

b. Suppose you are told that I have picked a face card. Now what is the probability it is a Queen?

c. Suppose instead you are told the card I have is black. Now what is the probability it is a Queen?

d. What do the above tell you about the random events "pick a Queen", "pick a face card" and "pick a black card"?

4 A certain binary information channel is more likely to transmit a 0 as an error for 1 than a 1 as an error for 0. The probability that a 1 is received in error is 0.1 (that is, $P(1 \text{ received} \mid 0 \text{ sent}) = 0.1$) and the probability that a 0 is received in error is 0.2. Write down the conditional probabilities for all four possible situations. If 1s and 0s are sent with equal probability, find the probability of receiving a zero. What is the probability that a zero was sent, given that a zero was received?

5 (For discussion, if you've not heard of it.) The famous and much debated Monty Hall problem is as follows: On a game show, a contestant is presented with three identical doors, behind one of which is a major prize, the others hiding a minor prize. The contestant gets to choose one door. If the contestant picks the door hiding a minor prize, then the host, Monty Hall, opens the door showing the other prize; if the contestant picks the door with the major prize, Monty randomly picks one of the doors hiding a minor prize and opens that. The contestant then has the choice of changing to the other door or not, and wins whatever is behind the door he or she has finally settled on.

The question is: should the contestant change to the other door? Can you prove your answer?

6 Suppose that we send a message of n bits in such a way that the probability of an error in any single bit is p and where the errors are assumed to be independent. Use the binomial probability distribution to write down the probability that:

a. k errors occur,

b. an even number of errors occur (including zero errors),

c. show that the answer in (b) can be expressed as

$$\frac{1 + (1 - 2p)^n}{2} \quad .$$

(*Hint: let $q = 1 - p$ and expand the expression* $\frac{(q+p)^n + (q-p)^n}{2}$.)

7 Taking $p = .001$ and $n = 100$ in Question 6, find the probability that

a. there is no error,

b. there is an undetected error when a simple parity check is used.

8 Check whether the following can be ISBNs, and if not, then assuming it is the check digit that is wrong, alter the check digit so that they are valid ISBNs:

$$0 - 552 - 08638 - X$$
$$0 - 576 - 08314 - 6$$

9 (A possible telephone number code)

Let $\mathcal{C}$ be the code of all 10 digit (decimal) numbers $\mathbf{x} = x_1 x_2 \cdots x_{10}$ that satisfy the two check equations

$$\sum_{i=1}^{10} x_i \equiv 0 \pmod{11}, \qquad \sum_{i=1}^{10} i x_i \equiv 0 \pmod{11}.$$

a. Estimate roughly how many such numbers $\mathbf{x}$ there are. (That is, estimate $|\mathcal{C}|$.)

b. Show that this code is single-error correcting and that it can also detect double errors caused by the transposition of two digits.

c. Correct the number $\mathbf{y} = 0680271385$.

10 (Introducing Chapter 4): A bridge deck is a set of 52 distinguishable cards. A bridge hand is any subset containing 13 cards and a bridge deal is any ordered partition of the bridge deck into 4 bridge hands.

a. Describe a simple but inefficient representation of an arbitrary bridge hand by assigning a 6-bit binary number to represent each card. How many bits are needed to describe a deal this way?

b. Show that no binary representation of an arbitrary bridge hand can use fewer than 40 bits. Give a representation using 52 bits.

c. Show that no representation of an arbitrary bridge deal can use fewer than 96 bits. Give a representation using 104 bits. This can be reduced to 101 bits with a little thought, can you see how?

Chapter 2: Error Correcting Codes

11 Using 9-character 8-bit even parity ASCII burst code, encode **HDforall** .

12 The integers from 0 to 15 inclusive are encoded as four-bit binary numbers with one parity check bit at the front, so that, for example, 4 is 10100 and 15 is 01111. A stream of such integers is then encoded into blocks of 4 with one check integer to give overall even parity, similar to the ASCII burst code.

- **a.** Show that the string of integers $1, 13, 2, 6$ has check number 8.
- **b.** The block 10010 11011 01110 00011 00110 is received. Find and correct the error (assuming at most one) and then decode.

13 A code $\mathcal{C}$ consists of all solutions $\mathbf{x} \in \mathbb{Z}_2^5$ (i.e., binary vectors of length 5) of the equation $H\mathbf{x}^T = \mathbf{0}$ where

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

- **a.** Find the codewords in $\mathcal{C}$.
- **b.** Find three linearly independent columns of H.
- **c.** Find three linearly dependent columns of H.

- **d.** Show that no two columns of H are linearly dependent.
- **e.** Show that no four columns of H are linearly independent.

14 For the Hamming (7,4)-code described in lectures
(a) encode 1010, (b) correct and decode 1001111.

15 Write down the parity check matrix of the Hamming (15,11)-code described in lectures, and hence correct and decode 001010101110110.

16 Using the Hamming (15,11)-code described in lectures:

- **a.** encode 10101010101,
- **b.** correct and decode 011100010111110.

17 What is the probability that a random sequence of $n = 2^m - 1$ '0's and '1's is a codeword in a Hamming code?

18 You receive a message which has been encoded as a codeword of the Hamming (7,4)-code. The probability of error in any single bit of the codeword while transmitted through a noisy channel is $p = .001$, and errors in different bits are independent. Find the probability that:

- **a.** there is no error,
- **b.** there are errors and they are correctly corrected.
- **c.** there are errors and they are not correctly corrected,
- **d.** when the code is used to detect but not correct errors (i.e. using a pure error detection strategy), there are undetected errors.

19 Suppose that linear code C has parity check matrix H.

- **a.** Prove that $d(C) = w(C)$.

b. Prove that $d = d(C)$ is the smallest integer r for which there are r linearly dependent columns in H modulo 2.

20 Suppose a binary linear code C has parity check matrix

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

a. Find its minimum distance d.

b. What are its error correction and detection capabilities?

c. Correct and decode the following received words (if possible), assuming that the first 3 bits are the information bits and we are using a single error correcting strategy.

(i) 010110

(ii) 010001

(iii) 100110

d. Find a standard form parity check matrix H' for the code C, and the corresponding generator matrix G'. Find a generator matrix G in similar fashion from the original parity check matrix H.

e. Explain how to extend the code C to give a new code $\widehat{C}$ with minimum distance $d+1$. (Prove that your new code has minimum distance $d+1$.) Write down a parity check matrix for $\widehat{C}$.

21 We wish to code the four directions N, S, E, W using a binary code.

a. Show that, if we require single error correction, then we need to use messages of at least 5 bits length. Find such a code of length 5.

b. Show that, if we require double error correction, then we need messages of at least 7 bits length. Construct such a code from messages of 8 bits.

22 Let $r \geq 3$ be an integer and suppose that $\mathbf{x} = x_1x_2\cdots x_n$ is a vector in $\mathbb{Z}_r^n$ (so $0 \leq x_i \leq r-1$ for all i).

a. For each nonnegative integer ρ, calculate the number of vectors in $\mathbb{Z}_r^n$ at distance ρ from $\mathbf{x}$.

b. Work out what the sphere-packing bound for t-error correcting radix r codes of length n should be, following the argument from the binary case.

c. Prove the radix r Hamming codes are perfect.

23 **a.** Construct the check matrix of a radix 5 Hamming code with parameters $m = 2$, $n = 6$, using the method given in lectures.

b. Using your check matrix, correct and decode the received word $\mathbf{y} = 410013$.

24 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

$$\begin{aligned} x_1 + x_2 + 3x_3 + 2x_4 &\equiv 0 \pmod 5 \\ x_1 + 2x_2 + 4x_3 + 3x_4 &\equiv 0 \pmod 5 \end{aligned}$$

a. Assuming that x_3 and x_4 are the information bits, find the codeword which encodes the message 21.

b. Which of the following is a valid codeword in C?

(1) 1122 (2) 1212 (3) 2323 (4) 4343

Chapter 3: Compression Codes

25 Decide whether the following codes are uniquely decodable, instantaneous or neither.

a. 0, 01, 11, 00 ;

b. 0, 01, 011, 111 ;

c. 0, 01, 001, 0010, 0011 ;

d. 00, 01, 10, 110, 111 .

26 Either prove the following code is uniquely decodable or find an ambiguous concatenated sequence of codewords:

$$\begin{array}{llll} \mathbf{c}_1 = 101, & \mathbf{c}_2 = 0011, & \mathbf{c}_3 = 1001, & \mathbf{c}_4 = 1110 \\ \mathbf{c}_5 = 00001, & \mathbf{c}_6 = 11001, & \mathbf{c}_7 = 11100, & \mathbf{c}_8 = 010100\,. \end{array}$$

(This is more difficult than Q25.)

27 Construct instantaneous codes, or show they cannot exist for the following:

a. radix 2, codeword lengths 1, 2, 3, 3, 3 ;

b. radix 2, codeword lengths 2, 2, 3, 3, 4, 4, 4 ;

c. radix 3, codeword lengths 1, 2, 3, 3, 3, 3, 3, 3, 3 ;

d. radix 3, codeword lengths 1, 1, 2, 2, 3, 3, 3, 3 .

Can any of these be shortened and if so how?

28 What is the minimum radix that would be needed to create a UD-code for the source $S = \{s_1, s_2, s_3, s_4, \dots, s_8\}$ with respective codeword lengths

a. $1, 1, 2, 2, 2, 3, 3, 4$

b. $2, 2, 2, 4, 4, 4, 4, 5$

29 Find binary Huffman codes and their expected codeword lengths for:

a. $p_1 = 1/2,\ p_2 = 1/3,\ p_3 = 1/6$;

b. $p_1 = 1/3,\ p_2 = 1/4,\ p_3 = 1/5,\ p_4 = 1/6,\ p_5 = 1/20$;

c. $p_1 = 1/2,\ p_2 = 1/4,\ p_3 = 1/8,\ p_4 = 1/16,\ p_5 = 1/16$;

d. $p_1 = 27/40,\ p_2 = 9/40,\ p_3 = 3/40,\ p_4 = 1/40$.

30 A random experiment has seven outcomes with corresponding probabilities

$$1/3,\ 1/3,\ 1/9,\ 1/9,\ 1/27,\ 1/27,\ 1/27.$$

The experiment is to be performed once and the outcome transmitted across the country. The telegraph company provides two services. Service 1 transmits binary digits at \$2.00 per digit and service 2 transmits ternary digits $\in \{0, 1, 2\}$ at \$3.25 per digit. You are to select a service and design a code to minimize expected cost.

a. Which service should be selected? What code should be used? What is the expected cost?

b. If the ternary cost is changed, at what new value of the cost would you change your mind?

***31** Prove that the (binary) Huffman code for a 2^n symbol source where each symbol has equal probability is a block code of length $n \geq 1$. (Hint: induction.)

***32** Suppose that we have an n symbol source where the ith symbol occurs with frequency f_i, where f_i is the ith Fibonacci number and $f_1 = f_2 = 1$. Describe the standard binary Huffman code for this source. (NOTE: $f_1 + f_2 + \cdots + f_n = f_{n+2} - 1$.)

33 Consider the alphabet $s_1, s_2, \cdots, s_8$ where the symbols occur with probabilities

0.22, 0.20, 0.18, 0.15, 0.10, 0.08, 0.05 and 0.02 respectively.

Code this source with a Huffman code of radix 4 using dummy symbols if necessary. What is the expected codeword length for this coding? Contrast it with the expected codeword length if another Huffman code is constructed by not introducing dummy symbols, but instead combining four symbols at a time as long as possible.

34 Consider the source $S = \{a, b\}$ with probabilities $p_1 = 3/4$ and $p_2 = 1/4$.

a. Find a binary Huffman code for the third extension S^3 of the source S. What is the average codeword length per (original) symbol.

b. Encode the message *aababaaaabaa* using this code.

35 Suppose we have two symbols which occur with probabilities $p_1 = 2/3$ and $p_2 = 1/3$. Consider the first, second and third extensions. Find Huffman codes for each extension and calculate the corresponding expected codeword lengths and expected codeword lengths per original symbol.

***36** Consider the Markov matrix

$$M = \begin{pmatrix} 1/3 & 1/4 & 1/4 \\ 1/3 & 1/2 & 1/4 \\ 1/3 & 1/4 & 1/2 \end{pmatrix}.$$

a. Show that M has eigenvalues 1, 1/4, 1/12 and find the equilibrium probabilities.

b. Explain why $\lim_{n\to\infty} M^n$ exists and find the limit. What do you notice about the answer?

37 A Markov source on symbols s_1, s_2, s_3 has transition matrix M and equilibrium vector $\mathbf{p}$ given as follows:

$$M = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.6 & 0.4 \\ 0.1 & 0.2 & 0.5 \end{pmatrix} \qquad \mathbf{p} = \frac{1}{17}\begin{pmatrix} 6 \\ 7 \\ 4 \end{pmatrix}.$$

a. Find Huffman codes for the equilibrium probability distribution and for the Markov source. Compare the expected codeword lengths in the two cases.

b. Encode the string of symbols $s_2s_2s_1s_1s_2s_3s_3$ using the Markov Huffman code.

c. Decode the code string 010001010 which was encoded using the Markov Huffman code.

38 A source has symbols $\{a, b, c, \bullet\}$ where $\bullet$ is the stop symbol. The probabilities of these symbols are $\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}$ respectively. Use arithmetic coding to encode the message $bac\bullet$ into a code number.

39 Three symbols s_1, s_2, s_3 and the stop symbol $s_4 = \bullet$ have probabilities $p_1 = 0.4$, $p_2 = 0.3$, $p_3 = 0.2$ and $p_4 = 0.1$.

a. Use arithmetic coding to encode the message $s_2s_1s_3s_1\bullet$.

b. Decode the codeword 0.12345 which was encoded using this arithmetic code.

40 Use the LZ78 algorithm to

a. encode "banana and bee" (including spaces).

b. decode

$$(0,t)(0,o)(0,\text{␣})(0,b)(0,e)(3,o)(0,r)(3,n)(2,t)(3,t)(2,\text{␣})(4,e)$$

Here "␣" denotes a space.

Chapter 4: Information Theory

41 A source S produces the symbols a, b, c, d, e with probabilities 1/3, 1/3, 1/9, 1/9, 1/9. Calculate the entropy of this source in bits.

42 Find the entropies (in appropriate units) for the sources in Questions 29, 33 and 35. Compare your answer with the expected codeword lengths of the Huffman codes you obtained previously.

Calculate the decimal entropy of the source in question 39, and compare to the length of the arithmetically coded message.

43 For the situation in Question 30, suppose the experiment was repeated many times and suitably coded (for long symbol words) before the outcomes were sent. What is the answer in part (a) and (b) now?

44 Find Shannon-Fano codes for the sources in Questions 29, 33 and 35.

45 A source S has 5 symbols $s_1, s_2, \ldots, s_5$ with probabilities $\frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{3}{20}, \frac{1}{10}$ respectively.

a. Calculate the entropy of S in **bits** to three significant figures.

b. Find a **ternary** Shannon-Fano code for S and its expected codeword length.

c. A **binary** Shannon-Fano code is constructed for S^4. Find the lengths of the two longest codewords in this code.

46 Let $H(p)$ be the entropy of a binary source with probability of emitting a 1 equal to p and probability of emitting a 0 equal to $1-p$. Show that on the interval $0 \leq p \leq 1$, the function $H(p)$ is nonnegative, concave down and has a unique maximum. Find this maximum and where it occurs and sketch the curve for $0 \leq p \leq 1$.

47 For the Markov sources with transition matrices as in Questions 36 and 37, find the Markov entropy H_M and equilibrium entropy H_E.

* **48** In a certain mathematics course, $\frac{3}{4}$ of the students pass, and the rest fail. Of those who pass, 10% own cars while half of the failing students own cars. All of the car owning students live at home, while 40% of those who do not own cars and fail, as well as 40% of those who do not own cars and pass, live at home.

a. Explain why the probability that a student lives at home or not given whether they own a car or not is independent of whether they have failed or not.

b. How much information (in bits) is conveyed about a student's result in the course if you know whether or not they own a car?

c. How much information (in bits) is conveyed about a student's result in the course if you know whether or not they live at home?

d. If a student's result, car owning status and place of residence are transmitted by three binary digits, how much of the total information in the transmission is conveyed by each digit? (Part (a) will be useful for the third digit.)

49 Consider a channel with source symbols $A = \{a_1, a_2\}$ and output symbols $B = \{b_1, b_2\}$, where $P(b_1 \mid a_1) = 0.8$ and $P(b_2 \mid a_2) = 0.6$. Suppose that $P(a_1) = 1/3$.

a. Using Bayes' Law, calculate $P(a_j \mid b_i)$ for all i, j.

b. Calculate the mutual information of the channel.

50 Find the channel capacity for the channel with source symbols $a_1 = 0$ and $a_2 = 1$, and received symbols $b_1 = 0, b_2 = 1$ and $b_3 = ?$, where $P(b_1|a_2) = P(b_2|a_1) = 0$, $P(b_3|a_1) = P(b_3|a_2) = q$ and $P(b_1|a_1) = P(b_2|a_2) = 1 - q$. (This is a special case of the Binary Symmetric Erasure Channel).

51 A channel has source symbols a_1, a_2, a_3 and received symbols b_1, b_2, b_3 and transition probabilities $P(b_i|a_i) = 1/2, P(b_j|a_i) = 1/4$ for all $i \neq j$. Find all the entropies, conditional entropies and the mutual information for the cases

a. $P(a_1) = P(a_2) = P(a_3) = 1/3$,

b. $P(a_1) = 1/2, P(a_2) = 1/3, P(a_3) = 1/6$,

c. What do you guess the channel capacity is and why?

***d.** Prove that your guess in (c) is correct ;)

52 Consider a channel with source symbols $A = \{a_1, a_2\}$ and output symbols $B = \{b_1, b_2, b_3, b_4\}$ where for some $0 \le x \le 1$ we have $P(a_1) = x$, $P(a_2) = 1 - x$ and

$$P(b_1 \mid a_1) = 5/7, \quad P(b_2 \mid a_1) = 2/7, \quad P(b_3 \mid a_2) = 1/10, \quad P(b_4 \mid a_2) = 9/10.$$

a. Calculate $H(A \mid B)$. How do you explain this result?

b. Hence find the capacity of the channel.

53 (For discussion.) Person A thinks of the name of a person who attends UNSW. Person B tries to determine who they are thinking of by asking A questions to which A has to reply simply "yes" or "no". Show that B can determine the name in 15 questions (assuming that they have access to the UNSW student database).

Chapter 5: Number Theory

54 Use the Euclidean Algorithm to find the gcd d of the following pairs a and b and express it in the form $d = xa + yb$ where $x, y \in \mathbb{Z}$:

(a) 324 and 3876, (b) 7412 and 1513, (c) 1024 and 2187.

55 List all the elements of $\mathbb{U}_{24}$ and hence evaluate $\phi(24)$.

Repeat for $\mathbb{U}_{36}$ and $\mathbb{U}_{17}$.

56 Find $\phi(72)$, $\phi(1224)$ and $\phi(561561)$.

57 Draw up addition and multiplication tables for $\mathbb{Z}_5$ and $\mathbb{Z}_6$ and explain why only the first one is a field.

58 Solve the congruence equations or show there is no solution:

(a) $6x \equiv 7 \pmod{17}$, (b) $6x \equiv 9 \pmod{12}$, (c) $11x \equiv 9 \pmod{13}$.

59 Find, if they exist, the inverse of 6 in

(a) $\mathbb{Z}_{11}$, (b) $\mathbb{Z}_{10}$, (c) $\mathbb{Z}_{23}$.

60 Use Euler's Theorem to find

(a) 2^{1001} in $\mathbb{Z}_{17}$, (b) the last two digits of 3^{1001}.

61 Find all the primitive elements for each of $\mathbb{Z}_{11}$ and $\mathbb{Z}_{17}$.

62 Use the polynomial Euclidean Algorithm to find the gcd $d(x)$ of the following pairs of polynomials $f(x)$ and $g(x)$ and express it in the form $d(x) = a(x)f(x) + b(x)g(x)$ where $a(x)$ and $b(x)$ are polynomials:

a. $x^3 + 1$ and $x^2 + 1$ in $\mathbb{Q}[x]$,

b. $x^3 + 1$ and $x^2 + 1$ in $\mathbb{Z}_2[x]$,

c. $x^2 - x + 1$ and $x^3 - x^2 - 1$ in $\mathbb{Z}_3[x]$.

63 Find

a. $x^5 + x^2 + 1 \pmod{x^2 + x + 1}$ in $\mathbb{Z}_2[x]$,

b. $x^5 + x^2 + 1 \pmod{x^2 + x + 1}$ in $\mathbb{Z}_3[x]$.

64 Write down addition and multiplication tables for both $\mathbb{Z}_2[x]/\langle x^2+x+1\rangle$ and $\mathbb{Z}_2[x]/\langle x^2+1\rangle$. Explain why only the first one is a field.

65 Which of the following are fields: $\mathbb{Z}_2[x]/\langle x^4 + x^2 + x + 1\rangle$, $\mathbb{Z}_3[x]/\langle x^4 + x^2 + x + 1\rangle$.

66 Construct the following finite fields, giving the table of powers of a primitive element γ (or α if α is primitive) and the corresponding linear combinations of the appropriate powers of the root α of the defining polynomial:

a. $\mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle$,

b. $\mathbb{Z}_2[x]/\langle x^4 + x^3 + x^2 + x + 1\rangle$,

c. $\mathbb{Z}_3[x]/\langle x^2 + x - 1\rangle$

Also list all the primitive elements in each field.

67 In $GF(16) = \mathbb{Z}_2(\alpha) = \mathbb{Z}_2[x]/\langle x^4 + x + 1\rangle$:

a. find the inverse of $\alpha^3 + \alpha + 1$,

b. evaluate $(\alpha^3 + \alpha + 1)(\alpha + 1)/(\alpha^3 + 1)$,

c. find the minimal polynomial of $\alpha^3 + \alpha + 1$

d. list all the minimal polynomials formed from powers of α.

68 List all the irreducible polynomials in $\mathbb{Z}_2[x]$ of degrees up to 4. Which of these are primitive in the fields they generate?

69 Show that 341 is a pseudo-prime to base 2 but not to base 3.

***70** Let a be an integer coprime to each of 3, 11 and 17.

a. Using Euler's theorem, show that $a^{560} \equiv 1$ modulo 3, 11 and 17.

b. Hence show that 561 is a Carmichael number.

71 Use Lucas' Test to test the primality of 97.

***72** Suppose n is a pseudoprime to base b and $\gcd(b-1, n) = 1$. Let $N = \dfrac{b^n - 1}{b - 1}$. Show that N is also a pseudoprime to base b and $\gcd(b-1, N) = 1$. Hence, show there are infinitely many composite pseudoprimes to any base $b > 1$.

73 Show that 561 is not a strong pseudo-prime base 2.

74 Let n be a fixed odd number and assume that

$$P(n \text{ passes } k \text{ Miller-Rabin tests } \mid n \text{ composite}) < 4^{-k}.$$

Use Bayes' rule to show that for large k we have approximately

$$P(n \text{ composite } \mid \text{ passes } k \text{ Miller-Rabin tests }) < 4^{-k}\frac{P(n \text{ composite})}{P(n \text{ prime})}.$$

75 Use Fermat factorization to factor 14647, 83411 and 200819.

Skip! 76 Use the Pollard-ρ method to factor $n = 8051$, using $f(x) = x^2 + 1$ and $x_0 = 1$. Repeat for $n = 201001$.

77 Let $n = 92131$. Factorise n with Fermat's method.

78 Let $N = 24497$.

a. Use Fermat factorisation to show that $N = 187 \times 131$.

b. Apply the Miller-Rabin test with $a = 2$ to give evidence that 131 is prime.

79 A number n is known to be of the form $n = (2^p - 1)(2^q - 1)$, where p, q, $2^p - 1$ and $2^q - 1$ are (unknown) primes with $p < q$. Find a method of factoring n in approximately p steps. Hence factorise 16646017.

Skip! 80 Consider the LFSR which implements the recurrence

$$x_{i+3} = x_{i+1} + x_i \quad \text{over} \quad \mathbb{Z}_2$$

Let $x_0 = 1$, $x_1 = 1$, $x_2 = 0$.

a. Generate the next 5 pseudo-random bits produced by the LFSR, namely $x_3, \ldots, x_7$.

b. What is the period of this LFSR?

Skip! 81 Trace the output of the pseudo-random number generators defined by

a. $x_{i+1} \equiv 7x_i + 1 \pmod{18}$, where $x_0 = 1$.

b. $x_{i+4} \equiv x_{i+3} + x_i \pmod{2}$, where $x_0 = 1$, $x_1 = x_2 = x_3 = 0$.

Chapter 6: Algebraic Coding

82 Set $m(x) = x^3 + x^2 + 1 \in \mathbb{Z}_2[x]$ and let α be a root of $m(x)$.

- **a.** Check that α is a primitive element of $GF(8)$.
- **b.** What is the minimal polynomial of α?
- **c.** A single error correcting BCH code is constructed over $GF(8)$ with primitive element α.
  - (i) What is the information rate of this code?
  - (ii) Encode [0,1,0,1].
  - (iii) Find the error and decode [1,0,1,1,0,1,1].

83 Set $p(x) = x^4 + x^3 + 1 \in \mathbb{Z}_2[x]$ and let β be a root of $p(x)$.

- **a.** Show that β is a primitive element of $GF(16)$, with minimal polynomial $p(x)$.
- **b.** A single error correcting BCH code is constructed over $GF(16)$ with primitive element β.
  - (i) What is the information rate of this code?
  - (ii) Encode [1,0,0,0,0,1,1,1,0,0,1].
  - (iii) Find the error and decode [0,0,0,0,0,1,1,1,1,0,0,0,1,1,0].

84 Set $q(x) = x^4 + x^3 + x^2 + x + 1$ and $F = \mathbb{Z}_2[x]/\langle q(x) \rangle$ (i.e. $F = GF(16)$).

- **a.** Find a primitive element γ for F and its minimal polynomial.
- **b.** Construct a single error correcting BCH code over F using γ and use it to
  - (i) encode [1,0,1,0,0,1,1,1,0,0,1],
  - (ii) find the error and decode [0,0,0,1,0,1,1,1,1,0,0,0,1,[illegible],1]

85 If β is as in Question 83, find the minimal polynomial for β^3. Construct a double error correcting code over $GF(16)$ using primitive element β and hence

- **a.** Encode [1,0,1,1,0,1,1].
- **b.** Decode [1,1,1,0,1,1,0,0,0,1,1,0,0,0,1], correcting any errors.
- **c.** Decode [1,1,1,0,1,1,0,0,0,1,1,0,1,0,1], correcting any errors.
- **d.** Decode [1,1,0,0,1,0,0,0,0,0,1,1,0,0,1], correcting any errors.

86 A double error correcting BCH code is based on the field $F = \mathbb{Z}_2[x]/\langle x^4 + x + 1 \rangle$.

- **a.** Encode the message [1,0,1,1,0,1,1].
- **b.** Decode the message [0,1,1,1,1,0,0,0,1,1,0,1,0,0,1], correcting any errors.

87 A coder is sending out 15-bit words, which are the coefficients of a polynomial $C(x)$ in $\mathbb{Z}_2[x]$ of degree 14 where $C(\alpha) = C(\alpha^2) = C(\alpha^3) = 0$ with $\alpha^4 + \alpha + 1 = 0$.

You receive $D(x) = 1 + x + x^2 + x^4 + x^8 + x^9 + x^{11} + x^{14}$, and assume that at most two errors were made. What was $C(x)$?

Skip! 88 A triple error correcting BCH code is constructed over $GF(16)$ with primitive element β, where β is a root of the polynomial $p(x) = x^4 + x^3 + 1$ (as in Question 83).

- **a.** What is the information rate of this code?
- **b.** Decode, with corrections, the message [1,1,0,0,0,0,1,0,0,0,0,0,1,0,0,1].
- **c.** Decode, with corrections, the message [1,0,1,0,1,0,0,1,0,0,1,0,1,0,1].

89 List all the cyclotomic cosets for the field GF(25).

Chapter 7: Cryptography

90 A message of 30 letters $m_1 m_2 \cdots m_{30}$ has been enciphered by writing the message into the columns of a 6×5 array in the order

$$\begin{pmatrix} m_1 & m_7 & \cdots & m_{25} \\ m_2 & m_8 & \cdots & m_{26} \\ \vdots & \vdots & \ddots & \vdots \\ m_6 & m_{12} & \cdots & m_{30} \end{pmatrix},$$

then permuting the columns of the array and sending the ciphertext by rows (reading across the first row, then the second row and so on). The resulting ciphertext is

`FSOTU OHFOI UIJNP RPUTM TELHE HQYEN`

a. Decipher the message.

b. Write down the permutation σ of the columns which was used to *encipher* the message (that is, column i becomes column $\sigma(i)$ for $i = 1, 2, \ldots, 5$).

c. Using the same permutation and the same method, encipher the message

`SELL ALL OIL SHARES BEFORE TAKEOVER`

(remove the spaces before enciphering).

91 The following has been enciphered using a simple transposition cipher with blocks of length 5. Decipher the message.

`HSTII AOSTN EAHHR ILTEC NEOCS ECROT EPDQS`

92 Consider the message

`ELIMINATE THE PERIODIC`

a. Encipher the message using plaintext feedback with the keyword `FISH`. (Remove spaces before enciphering.)

b. Encipher the same message with the same keyword, this time using ciphertext feedback (again removing spaces before enciphering).

c. The following ciphertext was enciphered using either plaintext feedback or ciphertext feedback with the keyword `FRED`.

`IZHB RCJC HNNI KBNL KITZ PEEK`

Determine which method was used for enciphering and decipher the message.

93 The following has been enciphered using a Vigenère cipher. It is suspected that the keyword has length 2.

```
NZYN  CYYF  YJYU  CHBW  LMMW  MSMW  LAYK  IXWS  YKUJ  WAJZ
YJMT  UKYV  IFNZ  YDYL  NWLK  IXUC  YQQG  LVZG  LFYS  LDSL
BJYW  BMHV  LWXQ  YSLK  NZYN  CYYF  YJYU  CHBW  LOUK  WGHK
CVYJ  YVOF  VJYS  ESVD
```

a. Calculate the index of coincidence and estimate the keyword length.

b. Find other evidence for a keyword length of 2.

c. Decipher the message, assuming that the keyword has length 2.

The following data may be useful:

letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
freq1	0	4	5	0	1	0	0	2	3	1	0	9	4	4	1	0	1	0	1	0	4	2	3	1	21	1
freq2	2	0	1	3	0	5	3	2	0	7	7	2	2	2	1	0	2	0	6	1	2	5	7	2	2	4

In the following questions, the standard encoding of the letters of the alphabet is used (see below). No letter is coded to 0 or 1 as they do not change under some of the arithmetic used. A space is coded as 2 and then the letters in order. The resulting numbers are then treated base 29.

0	1	␣	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28

***94** The following is an example of a matrix method of enciphering.

Given a message $\mathbf{m} = m_1m_2m_3m_4\ldots$ and a 2×2 matrix A, the message is enciphered by coding the letters as above and then using the transformation

$$A\begin{pmatrix} m_1 & m_3 & \ldots \\ m_2 & m_4 & \ldots \end{pmatrix} = \begin{pmatrix} c_1 & c_3 & \ldots \\ c_2 & c_4 & \ldots \end{pmatrix} \pmod{29}$$

to give the ciphertext $\mathbf{c} = c_1c_2c_3c_4\ldots$.

For example, given the matrix

$$A = \begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}$$

the message **m**=`NOANSWER` is enciphered as **c**=`WNWB1ZND` via

$$\begin{pmatrix} 2 & 3 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 16 & 3 & 21 & 7 \\ 17 & 16 & 25 & 20 \end{pmatrix} = \begin{pmatrix} 25 & 25 & 1 & 16 \\ 16 & 4 & 28 & 6 \end{pmatrix}$$

a. Using the matrix A from above, work out how to decipher the ciphertext and then decipher **c**=`LUVBRU`.

b. Using the above scheme but with an unknown matrix A, I sent the ciphertext

$$\mathbf{c} = \texttt{ZBXWCPIAIZFUZ0}$$

(note: the last symbol is zero "0", not capital O). But I made the mistake of starting my message with "HELLO". Using this information, work out in theory how you would decipher this and then do the deciphering.

c. Suppose all we had were the ciphertext (i.e. no crib) in **b**. Would we expect a unique meaningful message if we tried all possible matrices?

95 Using an RSA scheme with $n = 551$ and $e = 55$:

a. Encipher the message `HI` using the coding for letters given above and then the RSA encryption;

b. Find the deciphering exponent d and decipher the ciphertext $302, 241$.

96 Using an RSA scheme with $n = 391$ and $e = 235$, decipher the ciphertext

$$366,\ 14,\ 126,\ 126,\ 3,\ 249,\ 258,\ 126,\ 148,\ 30,\ 45,\ 366,\ 58,\ 30$$

where the letters have been coded as above.

97 Using an RSA scheme with $n = 1147$ and $e = 17$ we can encipher pairs of letters by encoding each as above and then writing the pair as a base 29 number (with a space at the end if there is only one letter) as in the examples

$$\texttt{OK} \to 17, 13 \to 17 \times 29 + 13 = 506 \to 506^{17} \equiv 410 \pmod{1147}$$

$$\texttt{A\_} \to 3, 2 \to 3 \times 29 + 2 = 89 \to 89^{17} \equiv 883 \pmod{1147}.$$

- **a.** Encipher the message `HELLO_`.
- **b.** What is the deciphering exponent?

98 A spy has designed a ciphering scheme as follows. Using an RSA scheme, he encodes a 3 letter key by mapping the first letter to the number a_1 and the second letter to the number a_2 etc using the standard encoding and then replacing (a_1, a_2, a_3) by $29^2 a_1 + 29 a_2 + a_3$.

This key is then encrypted using RSA encryption with $n = 10033$ and encryption exponent 1787. The spy then sends the message consisting of the RSA encoded key and the actual message enciphered by plaintext feedback.

- **a.** Factor $n = 10033$ using Fermat factorisation.
- **b.** Hence find $\phi(n)$, and from this calculate the decryption exponent d for the RSA coding of the key.
- **c.** Now decipher the message

8695IRDBHQIPVPBVKBRQ

sent using this code.

Skip! **99** In a simplified DSS scheme, the universal primes are $q = 13$, $p = 53$ and we have $g = 15$ of order 13 in $\mathbb{Z}_{53}$.

- **a.** Find Alice's public key if $e = 7$.
- **b.** If Alice picks $x = 5$ and sends a message with hash value $h = 9$, what will her signature be?
- **c.** Bob receives a message with a hash value $h = 10$ and signature $(2, 4)$. Is it genuine?

100 Find the unicity distance for the following ciphers:

- **a.** Vigenère cipher using a random keyword of length r;
- **b.** Vigenère cipher using a keyword of length $2s$ made up of a vowel then a consonant then a vowel etc.
- **c.** Polyalphabetic cipher using a random keyword of length r (here each letter of the keyword corresponds to a Caesar cipher);
- **d.** Transposition of length 10;
- **e.** DES cipher;
- **f.** RSA cipher with a 512-bit key.

Answers

1 (c) 373 is prime, but $323 = 17 \times 19$. **3** In order $\frac{1}{13}, \frac{1}{3}, \frac{1}{13}$, independent.

7 (a) $.999^{100} = .9047921471$ (b) $(1 + .998^{100})/2 - .9047921471 = .0044912554$.

8 (a) Correct (b) $0 - 576 - 08314 - 3$. **9** (a) roughly 10^8 (c) 0610271385

11 HDforallSYN **12** $2, 11, 12, 3$

13 (a) $\mathcal{C} = \{00000,\ 10101,\ 01011,\ 11110\}$, (b) No column is the sum of the other two: cols 1,2,3. (c) e.g. cols 1, 3, 5. (d) No two columns are equal, and none are zero. (e) H has only three pivot columns, so H has a maximum of three linearly independent columns.

14 (a) 1011010 (b) correct to 0001111, decode as 0111

15 correct: 001010101010110 decode: 11011010110

16 (a) 101101001010101 (b) correct to 010100010111110, decode to 00000111110

17 $2^k/2^n = 2^{-m}$

18 (a) $(1-p)^7 = .99302$. (b) $7p(1-p)^6 = .0069581$. (c) $1 - .99302 - .0069581 = .00002093$. (d) $7p^3(1-p)^4 + 7p^4(1-p)^3 + p^7 = 6.979 \times 10^{-9}$

20 (a) $d = 3$ (b) corrects 1 error (c) (i) is correct, decode as 010 (ii) correct to 011001, decode as 011 (iii) has at least two errors, cannot decode

23 (a) $H = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 3 & 4 \end{pmatrix}$ (b) correct to 410213, decode to 0213.

24 (a) 0221 (b) words (2) and (4)

25 (a) Neither (b) Uniquely Decodable (c) Neither (d) Instantaneous

26 Not UD: $s_4s_8s_2s_3 = s_7s_1s_5s_6$

27 (a) and (d) have $K > 1$. (b) $\{00, 01, 100, 101, 1100, 1101, 1110\}$. Can shorten 1110 to 111 (c) $\{0, 10, 110, 111, 112, 120, 121, 122, 200\}$. Can shorten 200 to 2.

28 (a) 4, (b) 3

29 (a) $\{1, 00, 01\}$, $L = 1.5$ (b) $\{00, 01, 11, 100, 101\}$, $L = 2.217$ (c) $\{1, 01, 001, 0000, 0001\}$, $L = 1.875$ (d) $\{0, 10, 110, 111\}$, $L = 1.425$.

33 (a) With dummy symbols, $L = 1.47$; without dummy symbols, $L = 2$.

34 (a) For example $(111) \to 1$, $(112) \to 001$, $(121) \to 010$, $(211) \to 011$, $(122) \to 00000$, $(212) \to 00001$, $(221) \to 00010$, $(222) \to 00011$. The average codeword length is $79/96 \approx 0.823$. (b) with the given coding we get 001 010 1 011

35 $L(S) = 1$, $\frac{1}{2}L(S^2) = 0.9\dot{4}$, $\frac{1}{3}L(S^3) = 0.938$

36 (a) $\mathbf{p} = \frac{1}{11}(3, 4, 4)^T$. (b) Each column of the limit is $\mathbf{p}$.

37 (a) $\mathbf{p} = \frac{1}{17}(6, 7, 4)^T$; $L_M = 1.388$, $L_E = 1.588$. (b) 1010010111. (c) $s_3s_2s_2s_1s_2$.

38 Any number in $[0.4608, 0.4640)$, e.g. 0.461

39 (a) Any number in $[0.49264, 0.49360)$. The shortest is 0.493. (b) $s_1s_1s_3s_1s_3\bullet$.

40 (a) (0,b)(0, a)(0,n)(2, n)(2,␣)(4, d)(0,␣)(1,e)(0,e) (b) to be or not to be

41 $H(S) = 2.113$

42 For Q29: (a) $H(S) = 1.460$; (b) $H(S) = 2.140$; (c) $H(S) = 1.875$; (d) $H(S) = 1.280$. Q33: $H(S) = 1.377$ radix 4 units/symbol. Q35: use $H(S^n) = nH(S)$. Q39: $H(S) = 0.556$ decimal bits/symbol.

43 (a) Average cost of binary is \$4.58; Average cost of ternary is \$4.69. (b) \$3.17.

44 Q29: (a) lengths 1, 2, 3; (b)lengths 2, 2, 3, 3, 5; (c) lengths 1,2,3,4,4; (d) lengths 1,3,4,6.
Q33: radix 4 code, lengths 2,2,2,2,2,2,3,3.
Q35: S^1: lengths 1,2; S^2: lengths 2,3,3,4; S^3: lengths 2,3,3,3,4,4,4,5.

45 (a) 2.20; (b) 1.77; (c) 13 and 14

47 Q37: $H_M = 1.293$, $H_E = 1.548$. Q36: $H_M = 1.523$, $H_E = 1.573$.

48 (b) 0.12, (c) 0.03, (d) 0.811, 0.602, 0.777 respectively

49 (a) $P(a_1 \mid b_1) = 1/2$, $P(a_2 \mid b_1) = 1/2$, $P(a_1 \mid b_2) = 1/7$, $P(a_2 \mid b_2) = 6/7$. (b) $I(A, B) = 0.109$.

50 Channel capacity $1 - q$.

51 (a) $I(A, B) = 0.085$, $H(A, B) = 3.085$, $H(A|B) = 1.5$.
(b) $I(A, B) = 0.077$, $H(A, B) = 2.959$, $H(A|B) = 1.382$.

52 (a) $H(A|B) = 0$ as there is no uncertainty in the input once the output is known. (b) capacity is 1.

54 (a) $12 = 324 \times 12 - 3876$ (b) $17 = 1513 \times 49 - 7412 \times 10$
(c) $1 = 1024 \times 472 - 2187 \times 221$.

55 $\mathbb{U}_{24} = \{1, 5, 7, 11, 13, 17, 19, 23\}$, $\phi(24) = 8$. $\mathbb{U}_{36} = \{1, 5, 7, 11, 13, 17, 19, 23, 25, 29, 31, 35\}$, $\mathbb{U}_{17} = \{1, 2, \ldots 16\}$.

56 $\phi(72) = 24$, $\phi(1224) = 384$, $\phi(561561) = 253440$.

58 (a) $x \equiv 4 \pmod{17}$, (b) No solutions since $\gcd(6, 12) = 6$ is not a factor of 9, (c) $x \equiv 2 \pmod{13}$.

59 (a) $6^{-1} = 2$ in $\mathbb{Z}_{11}$, (b) no inverse in $\mathbb{Z}_{10}$, (c) $6^{-1} = 4$ in $\mathbb{Z}_{23}$. **60** (a) 2 (b) 03.

61 (a) $2, 6, 7, 8$ (b) $3, 5, 6, 7, 10, 11, 12, 14$.

62 (a) gcd $= 1$, $a(x) = \frac{1}{2}(1 + x)$, $b(x) = \frac{1}{2}(1 - x - x^2)$. (b) gcd $= x + 1$, $a(x) = 1$, $b(x) = x$. (c) gcd $= x + 1$, $a(x) = x$, $b(x) = -1$.

63 (a) 1 (b) $x + 2$. **65** Only the second one.

67 (a) $\alpha^2 + 1$ (b) $\alpha^3 + \alpha^2 + \alpha + 1$ (c) $x^4 + x^3 + 1$.

71 Bases 2, 3 do not give any information but base 5 does.

75 $14647 = 151 \times 97$, $83411 = 239 \times 349$ and $200819 = 409 \times 491$.

76 $8051 = 83 \times 97$, $201001 = 19 \times 71 \times 149$. **77** $n = 13 \times 19 \times 373$.

79 Use: $(n - 1)/2^p$ is odd. 127×131071 **80** (a) $0, 1, 0, 1, 1$ (b) 7

81 (a) cycle length 18. (b) cycle length 15.

82 (c) (i) 4/7, (ii) $[1, 0, 0, 0, 1, 0, 1]$, (iii) correct: $[1, 0, 1, 0, 0, 1, 1]$, decode: $[0, 0, 1, 1]$.

83 (b) (i) 11/15, (ii) $[1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1]$
(iii) correct to $[0, 0, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0]$, decode as $[0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0]$.

84 (a) E.g. $\gamma = \alpha + 1$, where $q(\alpha) = 0$. The min poly γ is $m(x) = x^4 + x^3 + 1$.
(b) Using the BCH code based on γ: (i) $[0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1]$,
(ii) correct to $[0, 0, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0]$, decode to $[0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0]$

86 (a)] $[0,1,1,0,1,1,0,1,1,0,1,1,0,1,1]$ (b) correct to $[0,1,1,1,1,0,0,0,1,0,0,1,1,0,1]$, decode to $[1,0,0,1,1,0,1]$

87 $C(x) = 1 + x + x^2 + x^4 + x^6 + x^8 + x^9 + x^{11} + x^{14}$.

88 (a) 1/3, (b) correct to $[1,1,0,0,0,0,1,0,1,0,0,1,1,0,1]$, decode to $[0,1,1,0,1]$ (c) correct to $[1,1,1,0,1,0,1,1,0,0,1,0,0,0,1]$, decode to $[1,0,0,0,1]$

90 (a) SHIP EQUIPMENT ON THE FOURTH OF JULY (b) The permutation is $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 1 & 3 \end{pmatrix}$, which can also be written as (12534) in cycle notation.

(c) FSKAL OEERO RLOEU ELVSL TAEBS ALREH

91 THIS IS ANOTHER ARTICLE ON SECRET CODES P Q

92 (a) JTAT MYIF MGHX TXYM DHZK (b) JTAT RGAM VZHQ KDYY YGGA (c) DID YOU GET THE RIGHT ANSWERS (plaintext feedback)

93 (a) index of coincidence $I_c = 0.0602$, estimated keyword length 1.253, suggests keyword length either 1 or 2.

(c) THE VIGENERE CIPHER USES A SERIES OF CAESAR CIPHERS BASED ON THE LETTERS OF A KEYWORD FOR NEARLY THREE HUNDRED YEARS THE VIGENERE CIPHER WAS CONSIDERED UNBREAKABL

94 (a) ATTACK (b) HELLO EVERYBODY **95** (a) 409, 182 (b) OK

96 MERRY CHRISTMAS **97** (a) 1037, 424, 991 (b) 953

98 (a) $n = 10033 = 127 \times 79$, (b) $\phi(n) = 78 \times 126 = 9828$, $d = 11$; key word is ARR, message I AM THE PIRATE KING

99 (a) 42, (b) $(5,1)$, (c) No

100 (a) $\lceil 1.47r \rceil$ (b) $\lceil 2.10s \rceil$ (c) $\lceil 1.47r \rceil$ (d) 7 (e) 18 (f) 160