Version JS/3

EGT3
ENGINEERING TRIPOS PART IIB

Thursday 6 May 2021 9.00 to 10.40

Module 4F5

ADVANCED INFORMATION THEORY & CODING

*Answer not more than **three** questions.*

All questions carry the same number of marks.

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*Write your candidate number **not** your name on the cover sheet and at the top of each answer sheet.*

STATIONERY REQUIREMENTS
Write on single-sided paper.

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed.
You are allowed access to the electronic version of the Engineering Data Books.

10 minutes reading time is allowed for this paper at the start of the exam.

The time taken for scanning/uploading answers is 15 minutes.

Your script is to be uploaded as a single consolidated pdf containing all answers.

1 (a) Do there exist integers a and b such that $a + b = 1099$ and $\gcd(a, b) = 11$? [10%]

(b) Compute $\gcd(2^{30} + 1, 2^{10} + 1)$, showing your working. [10%]

(c) Show that, for any integers a and b, $R_a((a-1)^{2b}) = 1$, and use this to find the remainder $R_{101}(123{,}000{,}000{,}123)$. [10%]

(d) Using the Chinese Remainder Theorem, compute the multiplicative inverse of 37 in arithmetic modulo $3 \times 4 \times 5 = 60$. [20%]

(e) Consider arithmetic over the Galois field GF(9) with multiplication defined modulo the irreducible polynomial $\pi(X) = 1 + X^2$.

(i) What is the multiplicative order of X? [10%]

(ii) What is the additive order of X? [10%]

(iii) Now consider arithmetic over GF(9) modulo the polynomial $\pi(X) = 1 + 2X + 2X^2$, for which X is now a generator of the multiplicative group. What is the order of X^4? [10%]

(iv) Now consider arithmetic over GF(121) and assume that α generates the multiplicative group. What is the value of α^{60}? [10%]

(f) Let $\mathbf{x}$ and $\mathbf{y}$ be semi-infinite sequences of linear complexities $\mathcal{L}(\mathbf{x})$ and $\mathcal{L}(\mathbf{y})$, respectively. Let $\mathbf{z}$ be the element-wise sum of the sequences $\mathbf{x}$ and $\mathbf{y}$. State an upper bound on the linear complexity $\mathcal{L}(\mathbf{z})$ and provide a justification for your bound. [10%]

2 (a) What are the possible lengths of a Reed-Solomon code defined over the Galois field GF(31)? [10%]

(b) A Reed-Solomon code over GF(31) is defined using $\alpha = 2$. What is its code length? [10%]

(c) Write out the Discrete Fourier Transform (DFT) matrix on which the code in part (b) is based. [10%]

(d) An information sequence of length 3 is mapped to a codeword by pre-pending the information sequence with two zeros and multiplying the resulting vector by the inverse DFT matrix. What is the code's parity-check matrix? [10%]

(e) The inverse DFT matrix for the code in (d) is

$$\mathbf{F}^{-1} = \begin{bmatrix} 25 & 25 & 25 & 25 & 25 \\ 25 & 28 & 14 & 7 & 19 \\ 25 & 14 & 19 & 28 & 7 \\ 25 & 7 & 28 & 19 & 14 \\ 25 & 19 & 7 & 14 & 28 \end{bmatrix}$$

What is the code's encoder matrix? [10%]

(f) A codeword is transmitted over a channel that makes at most one error over the length of a codeword, and the received word is $\mathbf{r} = [16, 16, 4, 14, 4]$, whose DFT is $\mathbf{R} = [23, 23, 18, 20, 27]$. What were the encoded information digits? [20%]

(g) Write the systematic encoder matrix for the Reed-Solomon code in part (f). [20%]

(h) A codeword is transmitted over an erasure channel and the received word is $\mathbf{r} = [?, ?, 7, 14, 28]$. What were the encoded information digits if the systematic encoder was used? [10%]

3 (a) Alice, Bob and Eve are users of a Trusted Authority (TA) server running a Rivest-Shamir-Adelmann (RSA) Public Key Cryptosystem (PKC). The TA has verified their identity. Alice, Bob and Eve have each picked two large random primes, computed their products m_A, m_B and m_E, picked public keys e_A, e_B and e_E, computed private keys d_A, d_B and d_E, and published the pairs (m_A, e_A), (m_B, e_B) and (m_E, e_E), respectively. Analyse each scenario below with respect to authenticity and privacy, specifying for each:

- who can read the message and how?
- who could have generated the message and how can this be verified?
- is this an intended scenario for an RSA PKC? If so, describe in a few words a real-world application where it can be used.

The scenarios are:

(i) Bob encodes a message into a number x between 0 and $m_A - 1$, and transmits $y = R_{m_A}(x^{e_A})$ over the public channel. [15%]

(ii) Eve encodes the message "Bob, this is Alice. Send £100 to account 012345." into a number x between 0 and $m_B - 1$ and transmits $y = R_{m_B}(x^{e_B})$ over the public channel. [10%]

(iii) Alice encodes the message "Bob, this is Alice. Send £100 to account 012345." into a number x between 0 and $m_A - 1$ and transmits $(x, y) = (x, R_{m_A}(x^{d_A}))$ over the public channel. [15%]

(iv) Alice encodes a message into a number x between 0 and $m_B - 1$ and transmits $(y_1, y_2) = (R_{m_B}(x^{e_B}), R_{m_A}(y_1^{d_A}))$ over the public channel. [10%]

(b) Suppose we are given i.i.d. observations $y^n = (y_1, \ldots, y_n)$ and we wish to decide whether they have been generated from distributions P_0 or P_1, respectively. We assume priors are unknown.

(i) Show that if y^n is distributed according to distribution Q, then

$$Q(y^n) = e^{-n\left(H(\hat{P}_{y^n}) + D(\hat{P}_{y^n} \| Q)\right)},$$

where $\hat{P}_{y^n}$ is the empirical distribution of the observation sequence y^n. [20%]

(ii) Using (b)(i), show that the likelihood ratio test can be expressed as

$$D(\hat{P}_{y^n} \| P_1) - D(\hat{P}_{y^n} \| P_0) \lessgtr \gamma$$

for some γ. [15%]

(iii) Calculate the Stein exponent assuming that P_0 and P_1 are the conditional distributions of the output of a binary-symmetric channel with crossover probability p when the inputs are 0 and 1, respectively. [15%]

4 Let $W^n(y^n|x^n)$ be the transition probability assignment for sequences of length n on a discrete channel. Consider the ensemble of codes in which M codewords are independently chosen, each with probability assignment $Q^n(x^n)$. Let the messages encoded into these codewords have a probability assignment q_m for $m = 1, \ldots, M$, and consider a maximum a-posteriori decoder, which, given y^n, chooses the message m that maximizes $q_m \cdot W^n(y^n|x_m^n)$, i.e.,

$$\hat{m} = \underset{m=1,\ldots,M}{\arg\max}\ q_m \cdot W^n(y^n|x_m^n).$$

Let $\bar{p}_e = \sum_{m=1}^{M} q_m \bar{p}_{e,m}$ be the average error probability over this ensemble of messages and codes.

(a) Show that

$$\bar{p}_e \leq \left(\sum_{m=1}^{M} q_m^{\frac{1}{1+\rho}} \right)^{1+\rho} \sum_{y^n} \left(\sum_{x^n} Q^n(x^n) W^n(y^n|x^n)^{\frac{1}{1+\rho}} \right)^{1+\rho} \tag{1}$$

[40%]

(b) Let the channel be memoryless with transition probabilties $W(y|x)$, let the letters of the codewords be independently chosen with probability $Q(x)$, and the messages be sequences of length n from a discrete memoryless source V defined over alphabet $\mathcal{V}$ with probability assignment $P_V(v)$, for $v = 1, \ldots, |\mathcal{V}|$. Show that (1) is equivalent to

$$\bar{p}_e \leq e^{-n(E_0(\rho,Q)-E_s(\rho))}, \tag{2}$$

where

$$E_s(\rho) \triangleq \log \left(\sum_{i=1}^{|\mathcal{V}|} P_V(v_i)^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

$$E_0(\rho) \triangleq -\log \sum_{y} \left(\sum_{x} Q(x) W(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho}.$$

[25%]

(c) Show that $\bar{p}_e \to 0$ as $n \to \infty$ if $H(V) < C$, where $H(V)$ is the source entropy and C is the channel capacity. [15%]

(d) Show that (2) is equivalent to

(i) the channel random-coding error exponent theorem when $P_V = \frac{1}{|\mathcal{V}|}$, [10%]

(ii) the source random-coding error exponent theorem when the channel is noiseless. [10%]

END OF PAPER

THIS PAGE IS BLANK

Numerical Answers

I don't believe in the value of numerical answers: they train the wrong skills. Students who use numerical answers tend to fiddle around until they get the right answer. If you are not getting the right answer on the first attempt, it is often indicative of the fact that you have misunderstood something fundamental and you should not just move on and find a way to get those numbers. Hence, I am not providing numerical answers for this exam. Full answers are in the crib, as usual.

If you disagree with my view, feel free to contact me and explain why I'm wrong and I'd be happy to change my mind and provide numerical answers.

Jossy Sayir