4F5 Advanced Information Theory and Coding 2026 Crib

Question 1

(a) (i) The first step of Euclid's algorithm consists of computing

$$\begin{aligned} R_{99}(2897643258) &= R_{99}(28 \times 100^4 + 97 \times 100^3 + 64 \times 100^2 + 32 \times 100 + 58) \\ &= R_{99}(28 + 97 + 64 + 32 + 58) = 81 \end{aligned}$$

where we used the fact that $R_{99}(100^d) = 1$ for all d. Therefore, we can determine that

$$\gcd(2897643258, 99) = \gcd(81, 99) = 9.$$

(ii) We compute the CRT representation of 2^{340}, i.e., $(R_{11}(2^{340}), R_{31}(2^{340}))$.

$$R_{11}(2^{340}) = R_{11}\left((2^{10})^{34}\right) = 1$$

where we used the fact that $R_{11}(a^{10}) = 1$ for all $0 < a < 11$ because the ordere of the multiplicative group of GF(11) is 10. We compute the second remainder as

$$R_{31}(2^{340}) = R_{31}\left((2^{5})^{68}\right) = 1$$

Since the CRT representation of 2^{340} is $(1, 1)$, we conclude that $R_{341}(2^{340}) = 1$.

(iii) $1 + X + X^4$ is potentially either divisible by $(1 + X)$ or the product of two quadratic terms. If the former, then we can write

$$(1 + X)(1 + aX + bX^2 + X^3) = 1 + (1 + a)X + (a + b)X^2 + (1 + b)X^3 + X^4)$$

which implies that $1 + a = 1 \Rightarrow a = 0 \Rightarrow (a + b) = 0 \Rightarrow b = 0$, and we end up with a cubic term since $1 + b = 1$, so $1 + X + X^4$ cannot be divisible by $(1 + X)$. Similarly, consider the product of two quadratic terms

$$(1 + aX + X^2)(1 + bX + X^2) = 1 + (a + b)X + abX^2 + (a + b)X^3 + X^4$$

and we note that any such product will have the same coefficient for x as for X^3, so there is no way to obtain $1 + X + X^4$ as a product of two quadratic polynomials. Hence, $1 + X + X^4$ is irreducible.

(iv) The order of $GF(128)$ is 127 which is a prime number, so the order of all elements (including X) in $GF(128)$ is 127 no matter which irreducible polynomial is used to define multiplication. Hence, any irreducible polynomial of degree 7 is primitive, so if $1 + X + X^7$ is irreducible as stated, it must be primitive.

(v) We compute $X^5 = 1 + X^2$, $X^6 = X + X^3$ and $X^7 = X^2 + X^4$, hence the companion matrix of X^3 is

$$\boldsymbol{X}^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

(b) (i) We pre-compute the powers of 2 of α,

$$\alpha^1 = 5,\ \alpha^2 = 25,\ \alpha^4 = 14,\ \alpha^8 = 8,\ \alpha^{16} = 17,\ \alpha^{32} = 7$$

and note that $y_A = \alpha^{x_A} = \alpha^{21} = \alpha^{16+4+1} = 17 \times 14 \times 5 = 15$.

(ii) We pre-compute the powers of 2 of y_B,

$$y_B^1 = 12,\ y_B^2 = 3,\ y_B^4 = 9,\ y_B^8 = 34,\ y_B^{16} = 28,\ y_B^{32} = 32$$

then compute the common key as $y_B^{x_A} = y_B^{21} = y_B^{16+4+1} = 28 \times 9 \times 12 = 16$.

(c) (i) Eve can compute $E_k = Y_k - X_k$ for $k_0 \leq k < k_0 + L$. To recover the coefficients, she can set up linear equations of the form $E_k = a_1 E_{k-1} + a_2 E_{k-2} + \ldots + a_\ell E_{k-\ell}$ for $k = k_0 + \ell, \ldots, k_0 + 2\ell - 1$, giving ℓ equations for ℓ unknown coefficients. Hence, if $L = 2\ell$, this gives Eve a sufficient window to recover the coefficients and the state at time $k_0 + \ell$. Working forward and backward from this known state with known coefficients, she can recover the whole pseudo-random sequence and hence substract it from the cyphertext to recover the whole plaintext.

(ii) Blahut's theorem states that the linear complexity in the time domain equals the Hamming weight in the frequency domain. Hence, by performing a DFT of length $q - 1$ of the pseudo-random sequence $E_{k_0}, \ldots, E_{k_0+q-2}$ and counting the number of non-zero DFT coefficients, we obtain ℓ.

Question 2

(a) For example GF(8), GF(29) and GF(64). Essentially, any answer GF(q) for which q is a power of a prime and $q - 1$ is divisible by 7 works.

(b) Any lengths that divide 60, i.e., $3, 4, 5, 6, 10, 12, 15, 20, 30, 60$. I did not include 1 and 2 although they also divide 60, because RS codes of lengths 1 and 2 are useless.

(c)

$$\boldsymbol{F} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 9 & 20 & 58 & 34 \\ 1 & 20 & 34 & 9 & 58 \\ 1 & 58 & 9 & 34 & 20 \\ 1 & 34 & 58 & 20 & 9 \end{bmatrix}$$

(d) The rate is $R = 3/5$ and the parity check matrix consists of the first two rows of the DFT matrix

$$\boldsymbol{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 9 & 20 & 58 & 34 \end{bmatrix}$$

(e) We compute the DFT of the received word $\boldsymbol{r}$,

$$\boldsymbol{R} = \boldsymbol{r}\boldsymbol{F} = (3, 41, 11, 39, 26)$$

We now observe that, since $\boldsymbol{R} = \boldsymbol{C} + \boldsymbol{E}$ and the first two components of $\boldsymbol{C}$ are zero, we have $(E_0, E_1) = (3, 41)$. To determine $41/3$ in $GF(61)$, we could use Euclid's extended gcd algorithm to determine $1/3$, but it's easier to note that $3 \times 20 = 60 = -1$ hence $3 \times 3 \times 20 \times 20 = (-1)^2 = 1$, so $1/3 = 3 \times 20^2 = 41$, hence, confusingly, $41/3 = 41^2 = 34$ and hence we recover the whole error sequence in the frequency domain using the recurrence relation $E_k = 34E_{k-1}$,

$$\boldsymbol{E} = (3, 41, 52, 60, 27)$$

and subtract it from the received sequence to obtain

$$\boldsymbol{C} = (0, 0, 20, 40, 60)$$

giving the information sequence $(20, 40, 60)$.

(f) We replace the erasures by zeros and compute the DFT of the resulting "received" word

$$\boldsymbol{R} = (50, 0, 0, 33, 11)\boldsymbol{F} = (33, 20, 9, 50, 16)$$

We now recall that the erasures were at positions 1 and 2 (indexing from 0) so construct the error locator polynomial in canonical form

$$c(D) = (1 - \alpha^1 D)(1 - \alpha^2 D) = (1 - 9D)(1 - 20D) = 1 - 29D + 58D^2 = 1 - 29D - 3D^2.$$

This gives the recurrence relation $E_k - 29E_{k-1} - 3E_{k-2} = 0$ or $E_k = 29E_{k-1} + 3E_{k-2}$, allowing us to reconstruct the full error sequence in the frequency domain from its first two components $(E_0, E_1) = (33, 30)$,

$$\boldsymbol{E} = (33, 20, 8, 48, 13)$$

and, subtracting from he reveived sequence, we obtain

$$\boldsymbol{C} = (0, 0, 1, 2, 3)$$

giving the information sequence $(1, 2, 3)$.

(g) This requires a systematic encoder. We can obtain one by transforming the parity-check matrix to have the 2×2 identity matrix on the right (using the hint to determine $1/14 = 48$ in GF(61))

$$\boldsymbol{H}_{\text{sys}} = \begin{bmatrix} 52 & 32 & 35 & 1 & 0 \\ 10 & 30 & 27 & 0 & 1 \end{bmatrix}$$

and use the systematic matrix identities in the data book to determine

$$\boldsymbol{G}_{\text{sys}} = \begin{bmatrix} 1 & 0 & 0 & -52 & -10 \\ 0 & 1 & 0 & -32 & -30 \\ 0 & 0 & 1 & -35 & -27 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 9 & 51 \\ 0 & 1 & 0 & 29 & 31 \\ 0 & 0 & 1 & 26 & 34 \end{bmatrix}$$

Question 3

(a) (i) The random-coding error exponent $e_r(R)$ characterizes the exponential rate at which the ensemble average error probability over randomly generated codes decays with the source sequence length when compressing at rate R.

(ii) By definition, H satisfies

$$\max_{0\le\rho\le 1} \rho H - E_s(\rho) = 0.$$

This implies for any $0 \le \rho \le 1$,

$$\rho H - E_s(\rho) \le 0,$$

or equivalently

$$H \le \frac{E_s(\rho)}{\rho} \quad \forall \rho \in [0,1].$$

Since the equality holds for some ρ, it follows that

$$H = \min_{0\le\rho\le 1} \frac{E_s(\rho)}{\rho}.$$

(b) (i) The random-coding error exponent $E_r(R)$ characterizes the exponential rate at which the ensemble average error probability over randomly generated codes decays with the codeword length when communicating at rate R.

(ii) By definition, C satisfies

$$\max_{0\le\rho\le 1} E_0(\rho) - \rho C = 0.$$

This implies for any $0 \le \rho \le 1$,

$$E_0(\rho) - \rho C \le 0,$$

or equivalently

$$C \ge \frac{E_0(\rho)}{\rho} \quad \forall \rho \in [0,1].$$

Since the equality holds for some ρ, it follows that

$$C = \max_{0\le\rho\le 1} \frac{E_0(\rho)}{\rho}.$$

(c) (i) The function $E_0(\rho)$ for the erasure channel is

$$\begin{aligned} E_0(\rho) &= -\log \sum_{y\in\{0,1,?\}} \left(\sum_{x\in\{0,1\}} Q(x) W(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} \\ &= -\log \left(2^{-(1+\rho)}(1-\delta) + 2^{-(1+\rho)}(1-\delta) + \delta \right) \\ &= -\log(\delta + 2^{-\rho}(1-\delta)) \end{aligned}$$

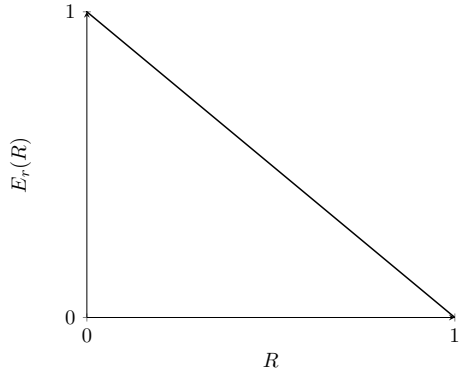
(ii)

$$C = \lim_{\rho \to 0} \frac{E_0(\rho)}{\rho} = \frac{0}{0}.$$

Since $E_0(0) = 0$, this is a 0/0 form, and by L'Hôpital's rule,

$$C = \lim_{\rho \to 0} \frac{E_0(\rho)}{\rho} = \frac{dE_0(\rho)}{d\rho}\Big|_{\rho=0} = I(X;Y) = 1 - \delta.$$

(iii) The error exponent is $E_r(\rho) = \max_{\rho \in [0,1]} (1-R)\rho = 1 - R$ and is sketched below



(iv) The random coding error exponent characterizes the exponential decay rate of the error probability averaged over the random code ensemble. Even when $\delta = 0$, the ensemble still contains poorly performing codes. As a result, the random coding error exponent remains finite, despite the existence of codes that achieve zero error probability.

Question 4

(a) From the lectures, the type 0,1 error probabilities are

$$\text{Type-0 error} : \epsilon_0\big(P_{X^n}, T\big) = \mathbb{P}\big[\text{test decides } h_1 \,|\, h_0 \text{ is true }\big] = P_{X^n}\big[T(X^n) = h_1\big]$$
$$\text{Type-1 error} : \epsilon_1\big(Q_{X^n}, T\big) = \mathbb{P}\big[\text{test decides } h_0 \,|\, h_1 \text{ is true }\big] = Q_{X^n}\big[T(X^n) = h_0\big].$$

(b) From the lectures, in the case the prior probabilities are unknown, the Neyman-Pearson setting considers the following trade-off

$$\alpha_\beta\big(P_{Y^n}, Q_{X^n}\big) = \min_{T : \epsilon_1(Q_{X^n}, T) \le \beta} \epsilon_0\big(P_{X^n}, T\big).$$

for $\beta \in (0,1)$.

(c) For every $\beta \in (0,1)$,

$$\lim_{n\to\infty} -\frac{1}{n} \log \alpha_\beta\big(P_{X^n}, Q_{X^n}\big) = D(Q\|P).$$

Similarly, for every $\alpha \in (0,1)$, if we define

$$\beta_\alpha\big(P_{X^n}, Q_{X^n}\big) \triangleq \min_{T:\ \varepsilon_0(P_{X^n}, T) \le \alpha} \varepsilon_1\big(Q_{X^n}, T\big),$$

then

$$\lim_{n\to\infty} -\frac{1}{n} \log \beta_\alpha\big(P_{X^n}, Q_{X^n}\big) = D(P\|Q).$$

(d) (i) Observe that

$$\begin{aligned} P_{X^n}\big[\phi_n(X^n) = \mathrm{e}\big] &= P_{X^n}\left[-\big(D(Q\|P) - \delta\big) \le \frac{1}{n}\sum_{i=1}^{n} \log \frac{P(X_i)}{Q(X_i)} \le \big(D(P\|Q) - \delta\big)\right] \\ &\le P_{X^n}\left[\frac{1}{n}\sum_{i=1}^{n} \log \frac{P(X_i)}{Q(X_i)} \le \big(D(P\|Q) - \delta\big)\right] \\ &\le P_{X^n}\left[\left|\frac{1}{n}\sum_{i=1}^{n} \log \frac{P(X_i)}{Q(X_i)} - D(P\|Q)\right| > \delta\right]. \end{aligned}$$

By the weak law of large numbers, the last term goes to zero as $n \to \infty$. The proof under Q_{X^n} is identical.

(ii)

$$\begin{aligned} P_{X^n}\big[\phi_n(X^n) = h_1\big] &= P_{X^n}\left[\frac{Q_{X^n}(X^n)}{P_{X^n}(X^n)} > e^{n[D(Q\|P)-\delta]}\right] \\ &\le e^{-n[D(Q\|P)-\delta]}\mathbb{E}_{P_{X^n}}\left[\frac{Q_{X^n}(X^n)}{P_{X^n}(X^n)}\right] \qquad \text{Markov's inequality} \\ &= e^{-n[D(Q\|P)-\delta]}\sum_{x^n}\frac{Q_{X^n}(x^n)}{P_{X^n}(x^n)}P_{X^n}(x^n) \\ &= e^{-n[D(Q\|P)-\delta]}. \end{aligned}$$

(iii) In classical binary hypothesis testing, the two errors must trade off. If we push one error down with its Stein exponent, then we cannot also make the other error go to 0 exponentially fast. Here, by allowing an erasure outcome, the error probabilities under P and under Q can both decay exponentially with exponents $D(Q\|P)$ and $D(P\|Q)$.