Solutions of the 4F5 Exam 2017

Jossy Sayir

23 May 2017

1. (a) Such a sequence is typical for all $\epsilon \geq 0$ because

$$P(x_1, \ldots, x_{100}) = p^{11}(1-p)^{89} = 2^{11\log p + 89\log(1-p)} = 2^{-nH_2(p)},$$

so the range is $\epsilon \in [0, \infty)$.

(b) $p^{10}(1-p)^{90} > p^{11}(1-p)^{89}$, hence it's the upper bound on $P(x_1, \ldots, x_{100})$ that restricts the value of ϵ.

$$2^{10\log p + 90\log(1-p)} \leq 2^{-n(H_2(p)-\epsilon)}$$
$$-10\log p - 90\log(1-p) \geq 100H_2(p) - 100\epsilon,$$

hence

$$\epsilon \geq H_2(p) + \frac{10}{100}\log p + \frac{90}{100}\log(1-p) = 0.03$$

The range is $\epsilon \in [0.03, \infty)$.

(c) $1 \geq \sum_{x_1 \ldots x_n \in A_{\epsilon,n}} P(x_1, \ldots, x_n) \geq |A_{\epsilon,n}| 2^{-n(H_2(p)+\epsilon)}$, hence $|A_{\epsilon,n}| \leq 2^{n(H_2(p)+\epsilon)}$

(d) $L_T = \lceil \log |A_{\epsilon,n}| \rceil \leq \lceil n(H_2(p)+\epsilon) \rceil = \lceil 100(H_2(0.11)+0.01) \rceil = \lceil 50.99 \rceil = 51$

(e) Since $2^{51} - |A_{\epsilon,n}| \geq 2^{51} - 2^{50.99} = 2^{50.99}(2^{0.01} - 1) \approx 10^{15} \times 10^{-3} \gg 1$, there are sequences of length 51 not used to encode typical sequences. Pick any such sequence $\hat{y}_1, \ldots, \hat{y}_{51}$ as an "escape" sequence. Encode $x_1, \ldots, x_{100}$ as follows

$$\begin{cases} \text{if } x_1, \ldots, x_{100} \in A_{\epsilon,n}, & \text{transmit codeword } y_1, \ldots, y_{51} \\ \text{if } x_1, \ldots, x_{100} \notin A_{\epsilon,n}, & \text{transmit } \hat{y}_1, \ldots, \hat{y}_{51} \text{ followed by } x_1, \ldots, x_{100}. \end{cases}$$

Since $E[L] = P(x_1, \ldots, x_n \in A_{\epsilon,n}) \cdot 51 + P(x_1, \ldots, x_n \not A_{\epsilon,n}) \cdot 151$, we have $E[L] \approx 51$ because the Asymtotic Equipartition Property (AEP) states that $P(x_1, \ldots, x_n \in A_{\epsilon,n}) \approx 1$.

(f) $|A_{\epsilon,n}| = \sum_{k=k_{\min}}^{k_{\max}} \binom{n}{k}$, where $k_{\min} = \min\{k : p^k(1-p)^{n-k} \leq 2^{-n(H_2(p)-\epsilon)}\}$ and $k_{\max} = \max\{k : p^k(1-p)^{n-k} \geq 2^{-n(H_2(p)+\epsilon)}\}$. All we need to do is find a p and an ϵ such that $k_{\min} = a$ and $k_{\max} = b$. For $k_{\max}$, we have

$$2^{k\log p + (n-k} \log(1-p) \geq 2^{-n(H_2(p)+\epsilon)}$$
$$-k\log p - (n-k)\log(1-p) \leq nH_2(p) + n\epsilon$$

hence

$$k \leq \frac{H_2(p) + \log(1-p) + \epsilon}{\log(1-p) - \log p} \cdot n = b$$

and similarly for $k_{\min}$, we have

$$k \geq \frac{H_2(p) + \log(1-p) - \epsilon}{\log(1-p) - \log p} \cdot n = a$$

We see that

$$\alpha = \frac{a+b}{2n} = \frac{H_2(p) + \log(1-p)}{\log(1-p) - \log p} = p$$

and

$$\beta = \frac{a-b}{2n} \log \frac{1-\alpha}{\alpha} = \frac{\epsilon}{\log(1-p) - \log p} \log \frac{1-p}{p} = \epsilon$$

hence

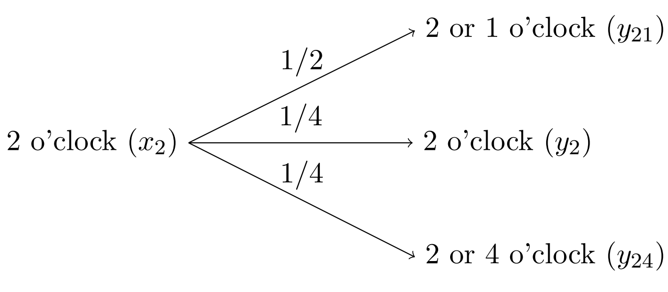
$$\sum_{k=a}^{b} = |A_{\epsilon,n}| \leq 2^{n(H_2(p)+\epsilon)} = 2^{n(H_2(\alpha)+\beta)}.$$

2. (a) There are 8 possible moves. If selected uniformly with $P_X(x) = \frac{1}{8}$ for all 8 possible x, we have $H(X) = 3$ bits per move.

   (b) There is a circular symmetry in the moves so we can concentrate on one move, say in the 2 o'clock direction. There are 4 possible output positions for this move:

   - east is unambiguously assigned to the 2 o'clock move
   - south can be attained from the 2 o'clock or from the 4 o'clock move
   - north and west can both be attained from the 2 o'clock or from the 1 o'clock move.

   By symmetry, for every input there will be a similar arrangement to the following channel for the 2 o'clock move:



   The channel is input-symmetric so capacity is achieved with a uniform input distribution $P_X(x) = 1/8$. For this, we have

$$\begin{cases} P_Y(y_2) = P_{Y|X}(y_2|x_2)P_X(x_2) = \frac{1}{4} \cdot \frac{1}{8} = \frac{1}{32} \\ P_Y(y_{21}) = P_{Y|X}(y_{21}|x_2)P_X(x_2) + P_{Y|X}(y_{21}|x_1)P_X(x_1) = \frac{1}{2} \cdot \frac{1}{8} + \frac{1}{2} \cdot \frac{1}{8} = \frac{1}{8} \\ P_Y(y_{24}) = P_{Y|X}(y_{24}|x_2)P_X(x_2) + P_{Y|X}(y_{24}|x_4)P_X(x_4) = \frac{1}{4} \cdot \frac{1}{8} + \frac{1}{4} \cdot \frac{1}{8} = \frac{1}{16}. \end{cases}$$

   There are 8 outputs of the type y_2 (unambiguously assigned to an input), 4 outputs of the type y_{21} (two ambiguously assigned knight positions), and 4 outputs of the type y_{24} (one ambigulously assigned knight position.) Hence we have

$$
\begin{aligned}
C = I(X;Y) &= H(X) - H(X|Y) = H(X) - \sum_y P_Y(y)H(X|Y=y) \\
&= H(X) - 8P_Y(y_2)H(X|Y=y_2) - 4P_Y(y_{21})H(X|Y=y_{21}) - 4P_Y(y_{24})H(X|Y=y_{24}) \\
&= 3 - 4P_Y(y_{21}) - 4P_Y(y_{24}) = 3 - \frac{4}{8} - \frac{4}{16} = \frac{9}{4} = 2.25 \text{ bits per move.}
\end{aligned}
$$

(c) The block length of a Reed-Solomon code over GF(8) must divide $8 - 1 = 7$. Since 7 is a prime number, the block length is $N = 7$.

(d) The probability of erasure is 3/4. Since we are told that $0 < K < N(1-p) = 7/4$ and K is an integer, the only solution is $K = 1$. The (7,1) Reed-Solomon code is essentially equivalent to a *repetition code*, for which the rate is $R = 1/7$ and the probability of a decoding failure is the probability that all code symbols are erased, i.e.,

$$P_e = \left(\frac{3}{4}\right)^7 = 0.1335.$$

3. (a) $R = 1 - \frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)dx} = 1 - \frac{\left[x^9/9\right]_0^1}{\left[x^2/24 + x^3/4 + x^4/24\right]_0^1} = 1 - \frac{1/9}{1/24 + 1//4 + 1/24} = \frac{2}{3}.$

(b) $\rho = \frac{C-R}{C} = \frac{1-p-2/3}{1-p} = \frac{1-.287-2/3}{1-.287} = 6.5\%$

(c) (i) $K = NR = 600$, hence the number of rows is $N - K = 300$.

(ii) The number of ones in the matrix is $M = 9(N-K) = 2700$, so the number of variable nodes of degree 3 is

$$\frac{\lambda_3}{3} \cdot M = \frac{3/4}{3} \cdot 2700 = 675.$$

(iii) The number of codewords is $2^K = 2^{600} \approx 10^{180}$.

(iv) No. The threshold only applies asymptotically as the block length goes to infinity. For a block length $N = 900$, there may well be patterns of $n < \theta N = 0.287 \cdot 900 = 258$ erasures that can not be recovered (this becomes less likely if $n \ll 258$.)

(d) (i) If the decoder for a code of block length N_k can decode n_k erasures for sure, then $d_k \geq n_k + 1$ because the optimal decoder can only be guaranteed to recover $d_k - 1$ erasures. Since the (suboptimal) iterative decoder is able to decode a proportion $\theta = 0.287$ of erasures with probability 1 as the block length goes to infinity, it follows that the code family must satisfy

$$\lim_{k\to\infty} \frac{d_k}{N_k} > \lim_{k\to\infty} \frac{n_k}{N_k} = \theta = 0.287$$

(ii) Only the optimal decoder is guaranteed to correct t errors where t is the largest number such that $2t < d_{\min}$. Hence the argument holds for the ML decoder but not for the sum-product decoder whose threshold could well be less than $\theta/2 = 0.1435$.

4. (a) (i) $E_s = \frac{1}{2}A^2 + \frac{1}{2}\kappa A^2 = \frac{A^2}{2}(1+\kappa^2)$. Note that this gives $A^2 = \frac{2E_s}{1+\kappa^2}$.

(ii) All symbols equally likely, the optimal MAP detector reduces to the minimum Euclidean distance rule. The decision regions are

$$\hat{X} = \begin{cases} \kappa A \text{ if } Y \geq \frac{\kappa A + A}{2} = \frac{A}{2}(1+\kappa) \\ A \text{ if } 0 \leq Y < \frac{A}{2}(1+\kappa) \\ -A \text{ if } -\frac{A}{2}(1+\kappa) \leq Y < 0 \\ -\kappa A \text{ if } Y < -\frac{A}{2}(1+\kappa) \end{cases}$$

(iii) Due to symmetry, it's sufficient to consider $X = A$ and $X = \kappa A$,

$$\begin{aligned} P_e &= \frac{1}{2}P(Y < \frac{A}{2}(1+\kappa)|X = \kappa A) + \frac{1}{2}\left[P(Y < 0|X = A) + P(Y \geq \frac{A}{2}(1+\kappa)|X = A)\right] \\ &= \frac{1}{2}P(N < \frac{A}{2}(1-\kappa)) + \frac{1}{2}\left[P(N < -A) + P(N > \frac{A}{2}(\kappa - 1))\right] \\ &= P(N > \frac{A}{2}(\kappa-1)) + \frac{1}{2}P(N < -A) \\ &= P\left(\frac{N}{\sqrt{N_0/2}} > \frac{A(\kappa-1)}{2\sqrt{N_0/2}}\right) + \frac{1}{2}P\left(\frac{N}{\sqrt{N_0/2}} > \frac{A}{\sqrt{N_0/2}}\right) \\ &= \mathcal{Q}\left(\sqrt{\frac{A^2(\kappa-1)^2}{2N_0}}\right) + \frac{1}{2}\mathcal{Q}\left(\sqrt{\frac{2A^2}{N_0}}\right) \\ &= \mathcal{Q}\left(\sqrt{\frac{E_s(\kappa-1)^2}{N_0(1+\kappa^2)}}\right) + \frac{1}{2}\mathcal{Q}\left(\sqrt{\frac{4E_s}{N_0(1+\kappa^2)}}\right) \end{aligned}$$

For $\kappa = 2$, the expression becomes

$$P_e = \mathcal{Q}\left(\sqrt{\frac{E_s}{5N_0}}\right) + \frac{1}{2}\mathcal{Q}\left(\sqrt{\frac{4E_s}{5N_0}}\right).$$

(b) (i) Multiplying by $h^\star/|h|$, we obtain $\tilde{Y} = |h|X + \tilde{N}$ where $\tilde{N} \sim \mathcal{CN}(0, N_0)$, so the effective signal is now $|h|X$, which, for $\kappa = 2$, takes on values in $\{-2|h|A, -|h|A, |h|A, 2|h|A\}$. Therefore, computing the error probability conditioned on h follows the same steps as in Part 4a, to yield

$$P_{e|h} = \mathcal{Q}\left(\sqrt{\frac{|h|^2E_s(\kappa-1)^2}{N_0(1+\kappa^2)}}\right) + \frac{1}{2}\mathcal{Q}\left(\sqrt{\frac{4|h|^2E_s}{N_0(1+\kappa^2)}}\right).$$

For $\kappa = 2$, the expression becomes

$$P_{e|h} = \mathcal{Q}\left(\sqrt{\frac{|h|^2E_s}{5N_0}}\right) + \frac{1}{2}\mathcal{Q}\left(\sqrt{\frac{4|h|^2E_s}{5N_0}}\right).$$

(ii) Using the approximation of the $\mathcal{Q}$ function, we obtain

$$P_{e|h} \approx \frac{1}{2}e^{-\frac{|h|^2 E_s}{10N_0}} + \frac{1}{4}e^{-\frac{2|h|^2 E_S}{5N_0}}$$

The probability of error averaged over h is

$$\begin{aligned} P_e &= \int_0^\infty \left[\frac{1}{2}e^{-\frac{E_s}{10N_0}x} + \frac{1}{4}e^{-\frac{2E_s}{5N_0}x}\right] e^{-x}dx \\ &= \frac{1}{2}\left(\frac{1}{1+\frac{E_S}{10N_0}}\right) + \frac{1}{4}\left(\frac{1}{1+\frac{2E_s}{5N_0}}\right) \end{aligned}$$

(iii) Using the same approximation of the $\mathcal{Q}$ function as we used for the fading channel, the error probability for the AWGN channel becomes

$$P_e \approx \frac{1}{2}e^{-\frac{E_s}{10N_0}} + \frac{1}{4}e^{-\frac{2E_s}{5N_0}}$$

which decreases exponentially in E_s/N_0, whereas the error probability for the fading channel decreases inverse proportionally to E_s/N_0.
However, this comparison is misleading because the two channels should really be compared on the basis of channel capacity rather than probability of error. Indeed, it is easy to improve the error probability of the fading channel with methods that exploit diversity, to transform the channel into equivalent AWGN channels of better error probability than our calculation for the fading channel suggested.

Q1 Information Theory and Typicality

25 attempts, Average mark 14.5 /20, Maximum 20, Minimum 8.

A popular question that was done fairly well and efficiently by most students, except part (f) that took much more work and was done well only by the top students. In part (d), we omitted to specify that $p = 0.11$ as in parts (a) and (b). Most candidates assumed $p=0.11$ despite our omission, but five candidates were affected by our omission and either assumed the worst case $p = 0.5$ in their upper bound or gave an answer in function of p instead of a numerical answer as required. All of these responses were counted as correct so the candidates' scores were not affected by our omission. A new version of the exam (JS/4) with the correct p specified is being supplied for the past exams archive so as not to confuse future candidates during their revision.

Q2 Information theory and Reed Solomon Coding

17 attempts, Average mark 14.1/20, Maximum 20, Minimum 9.

A slightly unusual question that required original thinking and hence was avoided by many students. Those who did the question found it mostly easy and did well. Several candidates got full points. Surprisingly, most of the candidates who attempted this question did well on the original/unusual "noisy chess channel" part of the question but had more difficulties in the standard Reed Solomon second part of the question.

Q3 LDPC coding

20 attempts, Average mark 10/20, Maximum 16, Minimum 4.

A fairly popular question for which the average performance was rather disappointing. Several sub-questions tested fundamental understanding of LDPC coding rather than material the students could have learned by heart, and anyone who really understands LDPC codes should have been able to answer most of these questions correctly in no time at all (e.g. the fact that the decoding algorithm is sub-optimal and hence that its performance is not determined by distance properties, or the fact that density evolution predicts performance only asymptotically in the block length). The poor average score is indicative of a weakness in teaching rather than a weakness of the cohort, as LDPC codes did not get enough coverage in 4F5. Students seem to have learned by rote but not gained fundamental understanding due to the rushed delivery of the material. This is addressed in the Part II reform of Information Engineering teaching as LDPC codes have been shifted from 4F5 to 3F7 and given a lot more time than was possible in 4F5, so we expect that 3F7 students would have done a lot better on this question.

Q4 Wireless Communication

25 attempts, Average mark 16.3/20, Maximum 19, Minimum 7.

A popular question that was done well by most students, partly because the question is very similar to questions asked on wireless communications in previous years. Students who had studied the methods well were able to do this question even if their understanding of the material was not perfect. The difference this year was that we considered an irregular constellation and most candidates were able to cope with this. Most candidates used the correct methods but many were let down by silly calculation mistakes and lost points when their responses were slightly wrong (losing a square, forgetting a factor 2, etc.)

Jossy Sayir (Principal Assessor)