Data Coding and Compression: Second Test

Master Program in Electrical and Computer Engineering, IST December 16, 2019

Name: ______________________________ Number: ________

Duration: 90 minutes. Part I scores: correct answer = 1 point; wrong answer = – 0.5 points.

Potentially useful facts: $e \simeq 2.72$; $\log_e 2 \simeq 0.69$; $\log_{10}(2) \simeq 0.30$; $\log_a b = \dfrac{\log_c b}{\log_c a}$; $\displaystyle\sum_{i=1}^{M} i = \frac{M(M+1)}{2}$.

Part I

1. Which of the following sequences is the result of applying Lempel-Ziv-Welch (LZW) coding to the sequence *"aaaaaa"* (assume that the alphabet is $\{a, b, c, d\}$ and that the dictionary indices start at 0)?

   a) *0,4,5*; ■
   b) *0,4,4,0*; □
   c) none of the above. □

2. Which of the following sequences results from applying a Lempel-Ziv-Welch (LZW) decoder to the sequence $0, 1, 2, 3, 4$ (assume that the alphabet is $\{a, b, c, d\}$ and that the dictionary indices start at 0)?

   a) *abcda*; □
   b) *abcdab*; ■
   c) none of the above. □

3. Consider an LZW decoder with alphabet $\{a, b, c, d\}$ and with the dictionary indices starting at 0. The decoding of the sequence "$1, 4, 5, ..., N$" produces a sequence of how many symbols?

   a) $(N-2)(N-1)/2$; ■
   b) $(N-1)N/2$; □
   c) none of the above. □

4. Let $X \in [0, \log_e 2]$ be a random variable (r.v.) with probability density function (p.d.f.) $f_X(x) = e^x$. Then,

   a) $h(X) < 0$; ■
   b) $h(X) = 0$; □
   c) $h(X) > 0$. □

5. Let $X \in \mathbb{R}$ be a r.v. with a Laplacian p.d.f. $f_X(x) = e^{-2\,|x|}$, which has variance $\sigma_X^2 = 1/2$. Then,

   a) $h(X) < \frac{1}{2}\log(\pi\, e)$; ■
   b) $h(X) = \frac{1}{2}\log(\pi\, e)$; □
   c) $h(X) > \frac{1}{2}\log(\pi\, e)$. □

6. Let $X \in [-1, 1]$ be a r.v. with uniform p.d.f. and $Y = |X|$. Then,

   a) $I(X;Y) = \infty$; ■
   b) $I(X;Y) = \log 2$; □
   c) none of the previous values. □

7. Let $X \in [0, 1]$ be a r.v. with uniform p.d.f. and $Y = -2X$. Then,

   a) $I(X;Y) = 0$; □
   b) $I(X;Y) = \infty$; ■
   c) none of the previous values. □

8. Let the r.v. $X \in [-1, 1]$ with p.d.f. $f_X(x) = \frac{1}{4}$, if $x \in [-1, 0]$, and $f_X(x) = \frac{3}{4}$, if $x \in [0, 1]$ be connected to a non-uniform quantizer Q with regions $R_0 = [-1,\, 0]$, $R_1 =]0,\, \frac{1}{3}]$, $R_2 =]\frac{1}{3},\, \frac{2}{3}]$, $R_3 =]\frac{2}{3},\, 1]$. The entropy of the output of the quantizer is

a) $H(Q(X)) < 2$ bits/symbol; ☐

b) $H(Q(X)) = 2$ bits/symbol;; ■

c) $H(Q(X)) > 2$ bits/symbol. ☐

9. Let $X \in [-\log_e(2), \log_e(2)]$ be a r.v. with p.d.f. $f_X(x) = \frac{1}{2} e^{|x|}$, connected to a 1-bit uniform quantizer. The optimal representative of the rightmost region R_1 is

a) $y_1 < \frac{1}{2} \log_e(2)$; ☐

b) $y_1 = \frac{1}{2} \log_e(2)$; ☐

c) $y_1 > \frac{1}{2} \log_e(2)$. ■

10. Let $X \in [-\log_e(2), \log_e(2)]$ be a r.v. with p.d.f. $f_X(x) = e^{-|x|}$, connected to a 1-bit uniform quantizer. The optimal representative of the rightmost region R_1 is

a) $y_1 < \frac{1}{2} \log_e(2)$; ■

b) $y_1 = \frac{1}{2} \log_e(2)$; ☐

c) $y_1 > \frac{1}{2} \log_e(2)$. ☐

11. Let X be the r.v. defined in question 10. Consider a 2-bit quantizer with the following representatives (codebook): $\mathcal{C} = \{y_0 = -\frac{1}{2},\, y_1 = -\frac{1}{4},\, y_2 = \frac{1}{4},\, y_3 = \frac{1}{2}\}$, and the following regions: $R_0 = [-\log_e(2), -a]$, $R_1 =]-a, 0]$, $R_2 =]0, a]$, $R_3 =]a,\, \log_e(2)]$. Which of the following choices of a leads to the lowest *mean squared error* (MSE)?

a) $a = \frac{1}{e}$; ☐

b) $a = \frac{1}{2} \log_e(2)$; ☐

c) $a = \frac{3}{8}$. ■

12. Let X, the r.v. defined in question 10, be connected to a 2-bit uniform quantizer. According to the high-resolution approximation, the mean squared error (MSE) satisfies

a) $\text{MSE} \simeq \frac{1}{24} \log_e(2)$; ☐

b) $\text{MSE} \simeq \frac{1}{48} \log_e(2)$; ☐

c) none of the previous answers. ■

13. Let $X \in [-1, 2]$ be a uniform r.v. connected to a 1-bit non-uniform quantizer with regions $R_0 = [-1, 0[$ and $R_1 = [0, 2]$. Then,

a) $\text{MSE} < 1/4$; ☐

b) $\text{MSE} = 1/4$; ■

c) $\text{MSE} > 1/4$. ☐

14. Let X be the r.v. defined in question 13, which, if connected to a 3-bit uniform quantizer, leads to a *signal-to-noise ratio* ($\text{SNR} = 10\, \log_{10}(\sigma_X^2/\text{MSE})$) of $\text{SNR} \simeq 18$ dB. For a 4-bit uniform quantizer,

a) $\text{SNR} \simeq 21$ dB; ☐

b) $\text{SNR} \simeq 24$ dB; ■

c) none of the previous answers. ☐

Part II

a) Let $X \in [-\log_e(2),\, \log_e(2)]$ be a r.v. with p.d.f. $f_X(x) = e^{-|x|}$, connected to a 2-bit non-uniform quantizer with the following regions: $R_0 = [-\log_e(2),\, -\log_e(\frac{4}{3})]$, $R_1 =]-\log_e(\frac{4}{3}),\, 0]$, $R_2 =]0,\, \log_e(\frac{4}{3})]$, and $R_3 =]\log_e(\frac{4}{3}),\, \log_e(2)]$. Compute the entropy of the output of the quantizer.

Advice: draw the density and the regions before answering the question.

b) Consider a r.v. $Y \in [-1,\, 1]$ with the following density: $f_Y(y) = \frac{1}{4}$, for $y \in [-1, 0]$ and $f_Y(y) = \frac{3}{4}$, for $y \in]0,\, 1]$. Let this r.v. be connected to a 1-bit non-uniform quantizer with regions $R_0 = [-1,\, \frac{1}{2}]$ and $R_1 =]\frac{1}{2},\, 1]$. Find the optimal representatives of these two regions. Is this an optimal quantizer?

Advice: draw the density and the regions before answering the question.

c) Consider a r.v. $Z \in [-1,\, 1]$ with uniform density connected to a R-bit uniform quantizer. What is the minimum value of R that guarantees that SNR ≥ 59 dB.