Norwegian University of Science and Technology
Department of Electronics and Telecommunications

TTT4110 Information and Signal Theory
Solution to the Exam of August 13, 2005

Problem 1 [14 points]

(a) [*2 points*]

We need to find a differential equation describing the relation between the input voltage $x(t)$ and the output voltage $y(t)$, both of which are defined in Fig. 1 of the problem statement.

Let $i(t)$ denote the current flowing through resistor R_1, and $i_1(t)$ and $i_2(t)$ respectively denote the currents flowing through capacitor C and resistor R_2, the flow directions of $i(t)$, $i_1(t)$, and $i_2(t)$ being chosen in such a way that

$$x(t) = R_1 i(t) + y(t),$$
$$y(t) = i_2(t) R_2,$$
$$i_1(t) = C \frac{dy(t)}{dt},$$

and

$$i(t) = i_1(t) + i_2(t).$$

We then see that

$$x(t) = R_1 \left[C \frac{dy(t)}{dt} + \frac{y(t)}{R_2} \right] + y(t),$$

or

$$R_1 C \frac{dy(t)}{dt} + \left(1 + \frac{R_1}{R_2} \right) y(t) = x(t), \tag{1}$$

which is the relation we were seeking.

(b) [*5 points*] Taking the Fourier transform of (1), we obtain

$$R_1 C \, j\Omega \, Y(\Omega) + \left(1 + \frac{R_1}{R_2} \right) Y(\Omega) = X(\Omega),$$

where $X(\Omega)$ and $Y(\Omega)$ respectively denote the Fourier transforms of $x(t)$ and $y(t)$. The frequency response $H(\Omega)$ of the filter with input $x(t)$ and output $y(t)$ is thus given by

$$H(\Omega) = \frac{Y(\Omega)}{X(\Omega)} = \frac{1}{R_1 C \, j\Omega + 1 + \frac{R_1}{R_2}}.$$

The magnitude and phase response of the filter in question can be respectively obtained by taking the magnitude and phase of $H(\Omega)$. This yields

$$|H(\Omega)| = \frac{1}{\sqrt{\left(1 + \frac{R_1}{R_2} \right)^2 + (R_1 C \, \Omega)^2}}$$

for the magnitude response, and

$$
\begin{aligned}
\angle H(\Omega) &= -\angle \left\{ \frac{1}{H(\Omega)} \right\} \\
&= -\angle \left\{ 1 + \frac{R_1}{R_2} + j\Omega R_1 C \right\} \\
&= - \arctan \frac{\Omega R_1 C}{1 + \frac{R_1}{R_2}}
\end{aligned}
$$

for the phase response, where the last equality follows from the fact that $1 + \frac{R_1}{R_2} > 0$ for all $R_1, R_2 > 0$.

It can be easily seen that the magnitude response $|H(\Omega)|$ is a monotonically decreasing function of Ω, and that $|H(0)| = \left(1 + \frac{R_1}{R_2}\right)^{-1}$. Therefore, the filter with transfer function $H(\Omega)$ attenuates more high frequency components than low frequency components, and is thus a low-pass filter.

(c) [*2 points*] The impulse response $h(t)$ of the filter can be obtained by taking the inverse Fourier transform of its transfer function $H(\Omega)$. In order to do so, remember first that

$$
\mathcal{F}\left\{ e^{-at} u(t) \right\} = \frac{1}{a + j\Omega},
$$

where $u(t)$ is the unit step function and $a \neq 0$. Bearing in mind that $H(\Omega)$ can be re-expressed as

$$
H(\Omega) = \frac{\frac{1}{R_1 C}}{\frac{R_1 + R_2}{R_1 R_2 C} + j\Omega} \triangleq \frac{\beta}{\alpha + j\Omega},
$$

we see that

$$
h(t) = \mathcal{F}^{-1}\{H(\Omega)\} = \beta e^{-\alpha t} u(t).
$$

(d) [*5 points*] To find the filter's response $y(t)$ when the input signal is

$$
x(t) = 10\cos(1000t) + \cos(3000t + \pi/4),
$$

we need to know its amplitude and phase responses at the angular frequencies $\Omega_1 = 1000$ rad/s and $\Omega_2 = 3000$ rad/s. We obtain, using the values $R_1 = 1\,k\Omega$, $R_2 = 10\,k\Omega$, and $C = 1\,\mu F$:

$$
|H(1000)| = \frac{1}{|1.1 + j|} \simeq 0.67,
$$

$$
|H(3000)| = \frac{1}{|1.1 + 3j|} \simeq 0.31,
$$

$$
\angle H(1000) = - \arctan \frac{1}{1.1} \approx -0.74 \text{ rad},
$$

$$
\angle H(3000) = - \arctan \frac{3}{1.1} \approx -1.22 \text{ rad}.
$$

The filter's response $y(t)$ to this signal is thus given by

$$
y(t) \approx 6.7\cos(1000t - 0.74) + 0.31\cos(3000t + \pi/4 - 1.22).
$$

2

Now, if the input signal to the filter is given by $x(t) = u(t)$, we can find the output $y(t)$ by using the equation

$$y(t) = h(t) * x(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\, d\tau.$$

We obtain, when $x(t) = u(t)$ and $h(t) = \beta e^{-\alpha t}u(t)$,

$$
\begin{aligned}
y(t) \quad &= \quad \int_{-\infty}^{\infty} \beta e^{-\alpha\tau}u(\tau)u(t-\tau)d\tau = \int_{0}^{\infty} \beta e^{-\alpha\tau}u(t-\tau)d\tau \\[2mm]
&= \begin{cases} 0 & \text{if } t < 0, \\[4mm] \int_{0}^{t} \beta e^{-\alpha\tau}\, d\tau = -\dfrac{\beta}{\alpha}\left[e^{-\alpha\tau}\right]_{0}^{t} = \dfrac{\beta}{\alpha}\left[1 - e^{-\alpha t}\right] & \text{if } t \geq 0, \end{cases}
\end{aligned}
$$

where $\frac{\beta}{\alpha} = \frac{R_2}{R_1+R_2} = \frac{10}{11}$, og $\alpha = \frac{R_1+R_2}{R_1 R_2 C} = 1100$.

Here we have used the fact that $u(\tau) = 0$ for $\tau < 0$, and similarly $u(t-\tau) = 0$ for $\tau > t$.

Problem 2 [13 points]

(a) [*3 points*]

   - The spectrum $X_a(F)$ of a signal $x_a(t)$ is not periodic. The signal $x_a(t)$ is therefore continuous.
   - The spectrum $X_a(F)$ shown in Fig. 2 of the problem statement is continuous. The signal $x_a(t)$ is therefore not periodic.
   - The spectrum $X_a(F)$ of $x_a(t)$ is a real and even function. The signal $x_a(t)$ is therefore also a real and even function.

(b) [*4 points*] We can avoid aliasing when sampling at a frequency F_s if the signal which is sampled is bandlimited to $F_s/2$. The signal $x_a(t)$ having a spectrum $X_a(F)$ which is bandlimited to $4\,\mathrm{kHz}$., the lowest frequency we can sample it at if we want to avoid aliasing (without using an antialiasing filter) is $F_{min} = 8\,\mathrm{kHz}$.

If a signal $x(n)$ is obtained by sampling an analog signal $x_a(t)$ with sampling frequency F_s, the relation between the spectrum $X(f)$ of $x(n)$ and the spectrum $X_a(F)$ of $x_a(t)$ is given by

$$X\left(\frac{F}{F_s}\right) = F_s \sum_{k=-\infty}^{\infty} X_a(F - kF_s).$$

Using also the relation $f = \frac{F}{F_s}$ (where f is the digital frequency), it is easy to draw the figure shown below.
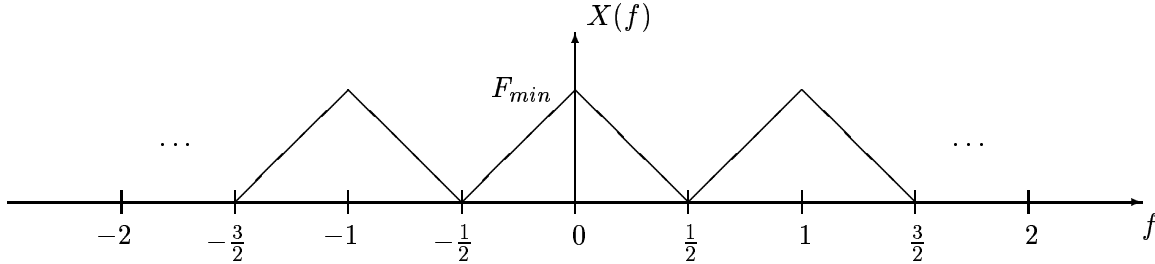

Figure 1: Spectrum $X(f)$ of the sampled signal $x(n)$

(c) [*6 points*] If we want to avoid aliasing when sampling the signal $x_a(t)$ at a frequency $F_s = \frac{3}{4}F_{min}$, all the frequency components of $x_a(t)$ which are above $\frac{F_s}{2} = \frac{3}{8}F_{min}$ must first be removed using an antialiasing filter. The magnitude response $|H(F)|$ of an ideal antialiasing filter which removes all the frequency components in $x_a(t)$ above $\frac{F_s}{2} = \frac{3}{8}F_{min} = 3\,\mathrm{kHz}$. while at the same time degrading $x(t)$ as little as possible is shown in Fig. 2.

The amplitude spectrum $|A(F)|$ of the signal at the output of this antialising filter (with input $x_a(t)$) is depicted in Fig. 3.

The amplitude spectrum $|X(f)|$ of the sampled signal $x(n)$ as a function of the digital frequency f when this antialiasing filter is used is shown in Fig. 4.
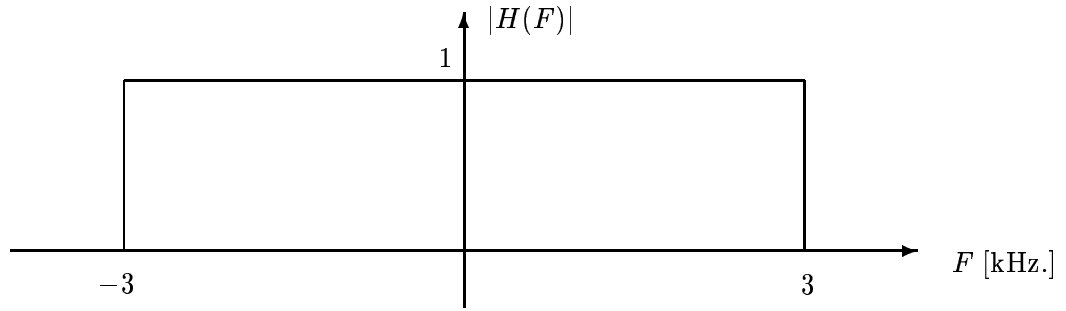
4

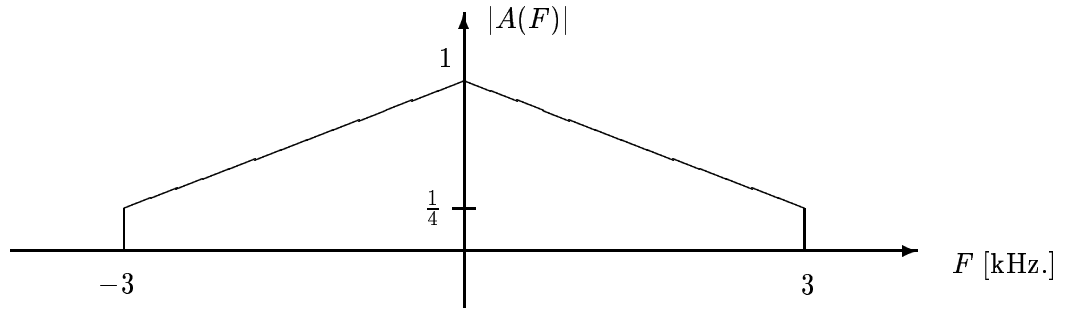Figure 2: Magnitude response $|H(F)|$

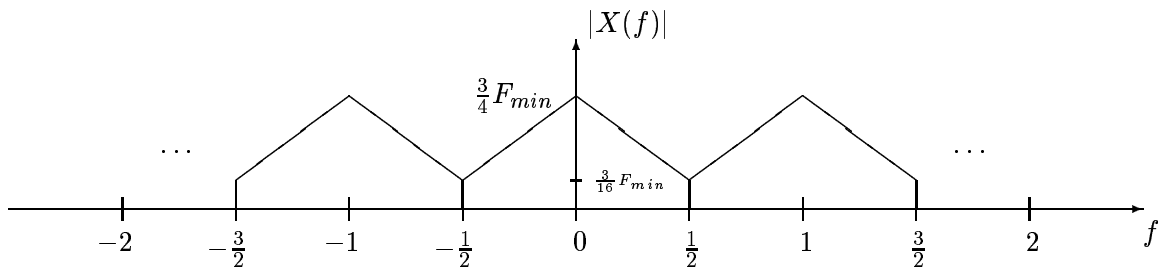Figure 3: Amplitude spectrum $|A(F)|$ of the signal at the output of the antialiasing filter

Figure 4: Amplitude spectrum $|X(f)|$ of the sampled signal $x(n)$ when the antialiasing filter is used.

Problem 3 [18 points]

(a) [*5 points*] $e(n)$ being a white noise process with variance $\sigma_E^2 = 4$, its autocorrelation function $R_{EE}(k)$ is given by

$$R_{EE}(k) = \sigma_E^2 \delta(k) = 4\delta(k).$$

The AR(1) process $x(n)$ is generated by sending white noise through a filter described by the difference equation

$$x(n) = ax(n-1) + e(n),$$

where $a = 0.9$.

Autocorrelation function is given by

$$
\begin{aligned}
R_{XX}(k) &= E[x(n)x(n-k)] \\
&= E[(ax(n-1) + e(n))x(n-k)] \\
&= aE[x(n-1)x(n-k)] + E[e(n)x(n-k)] \\
&= aR_{XX}(k-1) + E[e(n)x(n-k)].
\end{aligned}
$$

Assume first that $k > 0$. If we denote the unit step response of the filter by $h(n)$, and take into account that the filter is causal, we have that

$$x(n-k) = h(n-k) * e(n-k) = \sum_{l=0}^{\infty} h(l)\, e(n-k-l),$$

and thus

$$E\left[e(n)\, x(n-k)\right] = \sum_{l=0}^{\infty} h(l) E\left[e(n)\, e(n-k-l)\right] = 0,$$

where the last equality follows from the fact that white noise has uncorrelated samples. Autocorrelation function for $k > 0$ is therefore given by

$$R_{XX}(k) = aR_{XX}(k-1).$$

When $k = 0$, we have

$$
\begin{aligned}
R_{XX}(0) = E\left[x^2(n)\right] &= E\left[(ax(n-1) + e(n))^2\right] \\
&= a^2 E\left[x^2(n-1)\right] + E\left[e^2(n)\right] + 2aE\left[x(n-1)e(n)\right] \\
&= a^2 R_{XX}(0) + \sigma_E^2,
\end{aligned}
$$

where we have taken into account that $e(n)$ has zero mean. We thus obtain

$$R_{XX}(0) = \frac{\sigma_E^2}{1 - a^2}.$$

Remembering that $R_{XX}(-k) = R_{XX}(k)$ and combining the above equations, we obtain

$$R_{XX}(k) = R_{XX}(0)\, a^{|k|} = \frac{\sigma_E^2}{1 - a^2}\, a^{|k|} = 21.05 \cdot 0.9^{|k|}.$$
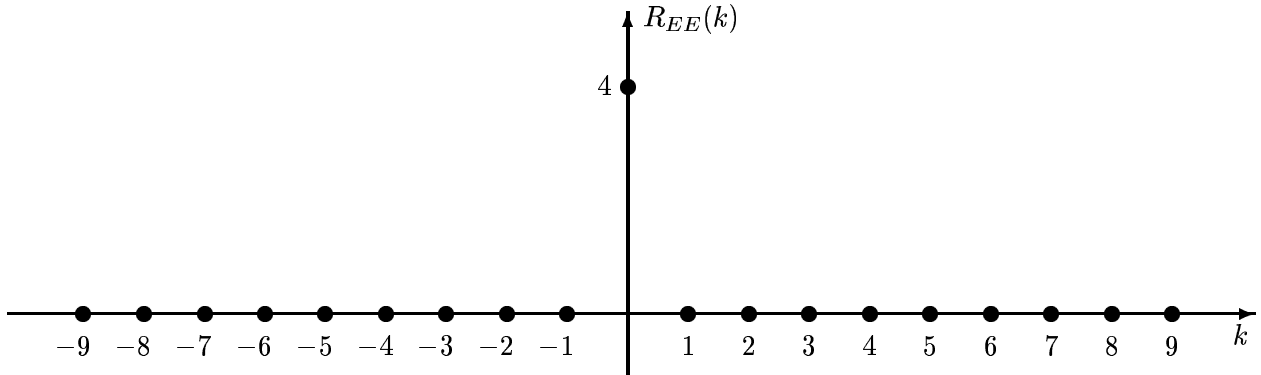
Figure 5: Autocorrelation function $R_{EE}(k)$ of the process $e(n)$
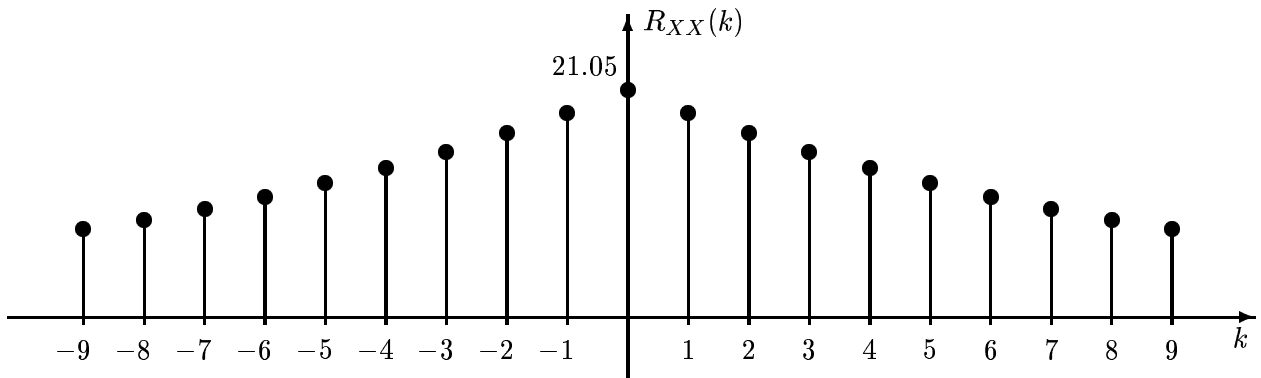

Figure 6: Autocorrelation function $R_{XX}(k)$ of the process $x(n)$

The autocorrelation functions $R_{EE}(k)$ and $R_{XX}(k)$ are respectively represented in Figs. 5 and 6.

The autocorrelation function $R_{ZZ}(k)$ of a stochastic process $z(n)$ evaluated at the value k tells us how much statistical dependence there is between two samples of $z(n)$ at distance k from each other. If $R_{ZZ}(k)$ is high then there is a strong statistical dependence between $z(n)$ and $z(n-k)$ (i.e. knowledge of $z(n-k)$ gives a lot of information about $z(n)$). Therefore, it is natural to expect that $R_{EE}(k) = 0$ when $k \neq 0$, because the samples of a white noise process are uncorrelated. However, the samples of the stochastic process $x(n)$ are related to each other through the relation

$$x(n) = 0.9x(n-1) + e(n),$$

and it is thus natural to expect a high correlation between samples of $x(n)$ that are close to each other, i.e. large values of $R_{XX}(k)$ for small values of k.

(b) [*2 points*] The power spectral density $S_{XX}(\omega)$ of the stochastic process $x(n)$ is given by

$$
\begin{aligned}
S_{XX}(\omega) &= \sum_{k=-\infty}^{\infty} R_{XX}(k)e^{-j\omega k} \\
&= \sum_{k=-\infty}^{0} R_{XX}(k)e^{-j\omega k} + \sum_{k=0}^{\infty} R_{XX}(k)e^{-j\omega k} - R_{XX}(0) \\
&= \sum_{k=0}^{\infty} R_{XX}(-k)e^{j\omega k} + \sum_{k=0}^{\infty} R_{XX}(k)e^{-j\omega k} - R_{XX}(0) \\
&= R_{XX}(0)\sum_{k=0}^{\infty} a^k e^{j\omega k} + R_{XX}(0)\sum_{k=0}^{\infty} a^k e^{-j\omega k} - R_{XX}(0) \\
&= R_{XX}(0)\left(\frac{1}{1-ae^{j\omega}} + \frac{1}{1-ae^{-j\omega}} - 1\right) \\
&= R_{XX}(0)\frac{1-a^2}{1-2a\cos\omega + a^2} = \frac{\sigma_E^2}{1-2a\cos\omega + a^2} \\
&= \frac{4}{1-1.8\cos\omega + 0.81},
\end{aligned}
$$

where we have used the fact that $a = 0.9 < 1$ to evaluate the geometric series, and the relation $\cos\omega = \frac{e^{j\omega}+e^{-j\omega}}{2}$.

(c) [*4 points*] The uniform quantiser $\hat{Q}$, with $N = 32$ quantisation levels, has to cover the interval $[-3\sigma, 3\sigma]$, where σ is the standard deviation of the signal input to the quantiser. The length Δ of a quantisation interval is thus given by

$$\Delta = \frac{6\,\sigma}{32} = \frac{3\,\sigma}{16}.$$

Let $q(n)$ denote the *quantisation noise*, i.e. the difference between the quantiser input signal and the quantiser output signal. It can be shown that $\sigma_Q^2 \approx \frac{\Delta^2}{12}$ for a uniform quantiser. Using this approximation, we see that when the signal $x(n)$ is input to $\hat{Q}$,

8

the quantisation noise variance becomes

$$\sigma_Q^2 \approx \frac{9\sigma_X^2}{256 \cdot 12} = \frac{9R_{XX}(0)}{256 \cdot 12} = \frac{9 \cdot 21.05}{256 \cdot 12} \approx 0.062,$$

where we har used the fact that $x(n)$ has zero mean, and thus $\sigma_X^2 = R_{XX}(0)$.

(d) [*5 points*] Knowing that $\hat{x}(n) = \alpha x(n-1)$, we need to find the value of α which minimises the variance of the prediction error $d(n) = x(n) - \hat{x}(n)$. Remembering that $x(n) = ax(n-1) + e(n)$, we have that

$$d(n) = (a - \alpha)x(n-1) + e(n),$$

and hence

$$\sigma_D^2 = E\left[((a-\alpha)x(n-1) + e(n))^2\right]$$
$$= (a-\alpha)^2\sigma_X^2 + \sigma_E^2,$$

since $e(n)$ and $x(n-1)$ are uncorrelated. We thus see that in order to minimise σ_D^2, we have to choose $\alpha = a$, i.e. $\alpha = 0.9$, and that the prediction error variance is then given by $\sigma_D^2 = \sigma_E^2$.

We now need to compute the variance σ_R^2 of the signal $r(n) = y(n) - x(n)$ when $\alpha = a$.

$$r(n) = y(n) - x(n)$$
$$= \hat{d}(n) + \hat{x}(n) - x(n)$$
$$= \hat{d}(n) - d(n).$$

We see that the reconstruction error equals the quantisation error of the quantiser $\hat{Q}$ with input $d(n)$. The reconstruction error variance is thus given by

$$\sigma_R^2 \approx \frac{9\sigma_D^2}{256 \cdot 12} = \frac{9\sigma_E^2}{256 \cdot 12} = \frac{9 \cdot 4}{256 \cdot 12} \approx 0.012,$$

since $\sigma_D^2 = \sigma_E^2$ when $\alpha = a$.

(e) [*2 points*] The noise variance introduced by DPCM is over 5 times smaller than when direct quantization is used. This was achieved by quantising the prediction error $d(n)$ instead of the signal $x(n)$ itself.

The gain was made possible by the fact that $x(n)$ is a highly correlated signal. The prediction error has therefore a lower variance than the signal $x(n)$ itself.

Problem 4 [10 points]

(a) [*2 points*] The quantity

$$I = \log_2 \left(\frac{1}{p} \right), \qquad \text{[bit]}$$

known as the *information content* of an event with probability p has the following properties:

- the more uncertain an event (the lower p), the larger its information content
- the information content of an event with probability one is zero, i.e. no information is gained by observing a certain event
- the information content is *additive*: the information gained by observing two independent events is the sum of the information contents of each one of the events

and is thus a good measure to quantify the amount of information gained by observing an event with probability p.

(b) [*3 points*] The entropy H of a discrete memoryless source producing symbols with probabilities $p_1, \ldots, p_N$ is the average amount of information that is gained by observing a symbol produced by the source, namely

$$H = E\{I\} = \sum_{i=1}^{N} p_i \log_2 \frac{1}{p_i} \quad \text{[bit]}$$

The entropy of a discrete memoryless source generating four symbols with probabilities $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}$, $p_3 = \frac{1}{8}$, and $p_4 = \frac{1}{8}$ is given by

$$H = \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \frac{1}{8} \log_2 8 + \frac{1}{8} \log_2 8 = 1.75 \text{ bits/source symbol.}$$

(c) [*5 points*] The code given in the problem statement is uniquely decodable because no codeword is the prefix of any other codeword.

The average number of bits per source symbol for this code is

$$\overline{L} = l_1 p_1 + l_2 p_2 + l_3 p_3 + l_4 p_4,$$

where l_i denotes the length of the codeword associated to the event with probability p_i. We thus obtain

$$\overline{L} = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8}$$
$$= 1.75 \text{ bits/source symbol.}$$

The lower limit for the average codeword length that is necessary to code a discrete memoryless source is given by its entropy. Since here $\overline{L} = H$, we see that there is no code with shorter average codeword length than that of the code suggested in the problem statement.