Assignment 8

Due time : 10:15, Nov 27, 2020 (Friday)

Name : 陈 稼 霖
Student ID : 45875852
Grade : ________

Problem 1 (2.11) Score: ________. Proof of the Kraft inequality for uniquely decodable codes.

(a) Assume a uniquely decodable code has lengths $l_1, \cdots l_M$. In order to show that $\sum_j 2^{-l_j} \leq 1$, demonstrate the following identity for each integer $n \geq 1$:

$$\left[\sum_{j=1}^{M} 2^{-l_j}\right]^n = \sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\cdots\sum_{j_n=1}^{M} 2^{-(l_{j_1}+l_{j_2}+\cdots l_{j_n})}$$

(b) Show that there is one term on the right for each concatenation of n codewords (i.e. for the encoding of one n-tuple $\boldsymbol{x}^n$) where $l_{j_1} + l_{j_2} + \cdots + l_{j_n}$ is the aggregate length of that concatenation.

(c) Let A_i be the number of concatenations which have overall length i and show that

$$\left[\sum_{j=1}^{M} 2^{-l_j}\right]^n = \sum_{i=1}^{nl_{max}} A_i 2^{-i}.$$

(d) Using the unique decodablility, upperbound each A_i and show that

$$\left[\sum_{j=1}^{M} 2^{-l_j}\right]^n \leq nl_{\max}.$$

(e) By taking the nth root and letting $n \to \infty$, demonstrate the Kraft inequality.

Proof: (a)

$$\left[\sum_{j=1}^{} M2^{-l_j}\right]^n = \left[\sum_{j_1=1}^{M} 2^{-l_{j_1}}\right]\left[\sum_{j_2=1}^{M} 2^{-l_{j_2}}\right]\cdots\left[\sum_{j_n=1}^{M} 2^{-l_{j_n}}\right] = \sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\cdots\sum_{j_n=1}^{M} 2^{-(l_{j_1}+l_{j_2}+\cdots+l_{j_n})}, \quad \forall n \geq 1. \tag{1}$$

(b) For each concatenation of n codewords $\boldsymbol{x}^n$, the length of the kth codeword is l_{j_k} and there is n codewords in total $(1 \leq k \leq n)$, so the aggregate length of the concatenation is $l_{j_1} + l_{j_2} + \cdots + l_{j_n}$. 一一对应

(c) Using the conclusion we obtained in (b), we have

$$\left[\sum_{j=1}^{M} 2^{-l_j}\right]^n = \sum_{\boldsymbol{x}^n} e^{-(\text{length of } \boldsymbol{x}^n)} = \sum_{i=1}^{nl_{\max}} A_i 2^{-i}. \tag{2}$$

(d) Because of unique decodablility (i.e. each concatenation should be different),

$$A_i \leq 2^i, \quad \forall i. \tag{3}$$

Thus,

$$\left[\sum_{j=1}^{M} 2^{-l_j}\right]^n = \sum_{i=1}^{nl_{\max}} A_i 2^{-i} \leq \sum_{i=1}^{nl_{\max}} 2^i 2^{-i} = \sum_{i=1}^{nl_{\max}} 1 = nl_{\max}. \tag{4}$$

2 x + 1

(e) Taking the nth root of equation (4), we have

$$\sum_{j=1}^{M} 2^{-l_j} \leq [nl_{\max}]^{1/n}. \tag{5}$$

Letting $n \to \infty$, we have

$$\sum_{j=1}^{M} 2^{-l_j} \leq \lim_{n\to\infty} [nl_{\max}]^{1/n} = 1, \tag{6}$$

which is the Kraft inequality. □

Problem 2 (2.12) Score: ________. A source with an alphabet size of $M = |\mathcal{X}| = 4$ has a symbol probabilities $\{1/3, 1/3, 2/9, 1/9\}$.

(a) Use the Huffman algorithm to find an optimal prefix-free code for this source.

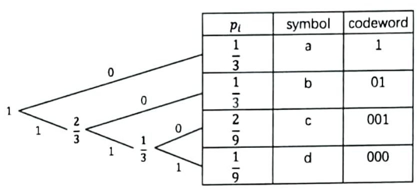
(b) Use the Huffman algorithm to find another optimal prefix-free code with a different set of lengths.

(c) Find another prefix-free code that is optimal but cannot result from using the Huffman algorithm.

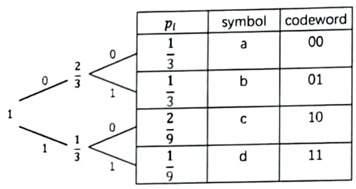
Solution: Suppose the four symbols corresponding to the probabilities $\{1/3, 1/3, 2/9, 1/9\}$ are a,b,c,d.

(a) An optimal prefix-free code for this source is shown in figure 1(a).

(b) Another optimal prefix-free code for this source is shown in figure 1(b).



(a) An optimal prefix-free code derived from the Huffman algorithm.



(b) Another optimal prefix-free code derived from the Huffman algorithm.

Figure 1: Two optimal prefix-free code scheme.

(c) Another prefix-free code that is optimal but cannot result from using the Huffman algorithm is shown in table 1.

Table 1: Another prefix-free code that is optimal but cannot result from using the Huffman algorithm

symbol	codeword
a	00
b	11
c	10
d	01

15

□

Problem 3 (2.14) Score: ________. Consider a source with M equiprobable symbols.

(a) Let $k = \lceil \log M \rceil$. Show that, for a Huffman code, the only possible codeword lengths are k and $k-1$.

(b) As a function of M, find how many codewords have length $k = \lceil \log M \rceil$. What is the expected codeword length $\bar{L}$ in bits per source code?

(c) Define $y = M/2^k$. Express $\bar{L} - \log M$ as a function of y. Find the maximum value of this function over $1/2 < y \le 1$. This illustrates that the entropy bound, $\bar{L} = H[X] + 1$, is rather loose in this equiprobable case.

①先论霍夫曼编码只能有两种长度的codeword

Lemma 2.5: if $p_i \ge p_j$, then $l_i \le l_j$. (利用反证法可得)

第一步将两个点加在一起，从而比其它点大，比其它点的codeword长1.

Solution: (a) For a Huffman code, if M is a power of 2, say $M = 2^k$, then the Huffman tree should be a full binary tree and the length is $k = \log_2 M$ for all codewords. If M is not a power of 2, say $M = 2^{k-1} + k_0$ where $0 < k_0 \le 2^{k-1}$, the Huffman tree should be a complete but not full binary tree. In this case, some codewords are at the bottom layer of the Huffman tree whose lengths are all $k = \lceil \log_2 M \rceil$, other codewords are at the bottom but one layer whose lengths are all $k - 1 = \lceil \log_2 M \rceil - 1$.

(b) Suppose the number of codewords with length k is x, then the number of codewords with length $k-1$ is $M - x$. The number of node at the bottom but one layer of the Huffman tree should be

$$\frac{x}{2} + (M - x) = 2^{k-1}, \tag{7}$$

so the number of codeword with length k is

$$x = 2M - 2^k. \tag{8}$$

The expected code length per source code is

$$\bar{L} = \frac{k(2M - 2^k) + (k-1)(2^k - M)}{M} = k + 1 - \frac{2^k}{M}. \tag{9}$$

②再论codeword中较长的那种长 $n = \lceil \log_2 M \rceil$.

利用Kraft不等式 $\sum 2^{-l_i} \le 1$

Huffman码的编码2叉树是完全2叉树，故取等 $\sum 2^{-l_i} = 1$.

设长度为n的有n_k个

则长度为$n-1$的有$M-n_k$个

$1 \le n_k \le M$

(当为满树时，右边取等 此时 $\log_2 M = n$)

15

$n_k \cdot 2^{-n} + (M - n_k) \cdot 2^{-(n-1)} = 1$.

故 $2^{n-1} < M \le 2^n$

$\Longrightarrow n = \lceil \log_2 M \rceil$

(c) Using $k = \log_2 \frac{M}{y}$, we have

$$\bar{L} - \log_2 M = -\log_2 y + 1 - \frac{1}{y}. \tag{10}$$

The derivative of the above function is

$$\frac{\mathrm{d}}{\mathrm{d}y} = -\frac{1}{y \ln 2} + \frac{1}{y^2} \begin{cases} > 0, & \frac{1}{2} \le y < \ln 2, \\ < 0, & \ln 2 < y < 1, \end{cases} \tag{11}$$

so the maximum value of the above function is

$$[\bar{L} - \log_2 M]_{\max} = [\bar{L} - \log_2 M]_{y = \ln 2} = -\log_2(\ln 2) + 1 - \frac{1}{\ln 2} = 0.086, \tag{12}$$

which means that

$$\bar{L} \le \log_2 M + 0.086 \le H(X) + 0.086. \tag{13}$$

Therefore, the entropy bound $\bar{L} = H[X] + 1$, is rather loose in this equiprobable case.

□

Problem 4 (2.21) Score: ________. A discrete memoryless source emits iid random symbols $X_1, X_2, \cdots$ Each random symbol X has the symbols $\{a, b, c\}$ with probabilities $\{0.5, 0.4, 0.1\}$, respectively.

(a) Find the expected length $\bar{L}_{\min}$ of the best variable-length prefix-free code for X.

(b) Find the expected length $\bar{L}_{\min,2}$, normalized to bits per symbol, of the best variable-length prefix-free code for X^2.

(c) Is it true that for any DMS, $\bar{L}_{\min} \ge \bar{L}_{\min,2}$? Explain your answer.

Solution: (a) The best variable prefix-free code for X is shown in figure 2(a), whose expected length is

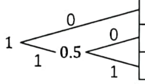
$$\bar{L}_{\min} = 0.5 \times 1 + 0.4 \times 2 + 0.1 \times 2 = 1.5 \quad . \tag{14}$$

(b) The best variable prefix-free code for X^2 is shown in figure 2(b), whose expected length normalized to bits per symbol is

$$\begin{aligned}\bar{L}_{\min,2} =& \frac{1}{2}(0.25\times 2+0.2\times 2+0.2\times 2+0.16\times 3+0.05\times 5+0.05\times 5+0.04\times 5+0.04\times 6+0.01\times 6)\\ =&1.39 \quad .\end{aligned} \tag{15}$$

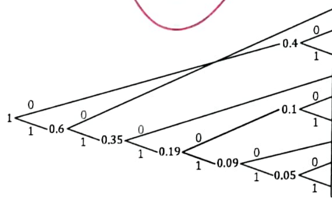
(c) It is true that for any DMS, $\bar{L}_{\min} \geq \bar{L}_{\min,2}$. One method for source coding of X^2 is to use the concatenation of two codewords of X as the codeword of X^2, whose expected length per symbol equals $L_{\min}$. Because this method is not necessarily the best coding method, $L_{\min}$ can not be less than $\bar{L}_{\min,2}$, which is the minimal expected length per symbol of X^2.

15



p_i	symbol X	codeword
0.5	a	0
0.4	b	10
0.1	c	11

(a) The best variable-length prefix-free code for X



p_i	symbol X^2	codeword
0.25	aa	10
0.2	ab	00
0.2	ba	01
0.16	bb	110
0.05	ac	11100
0.05	ca	11101
0.04	bc	11110
0.04	cb	111110
0.01	cc	111111

(b) The best variable-length prefix-free code for X^2

Figure 2: The best variable-length prefix-free code schemes.

□

Problem 5 (2.33) Score: ________. Perform an LZ77 parsing of the string 000111010010101100. Assume a window of length $W = 8$; the initial window is underlined above. You should parse the rest of the string using the Lempel-Ziv algorithm.

Solution: The rest of the string can be parsed as

$$\underline{00011101}\ \overbrace{001}^{u=7,n=3}\ \overbrace{0101}^{u=2,n=4}\ \overbrace{100}^{u=8,n=3}\ ,$$

whose corresponding encoded sequence is

$$011\ 111\ 00100\ 010\ 011\ 000.$$

5

□

Problem 6 (4.35 Aliasing) Score: ________. The following exercise is designed to illustrate the sampling of an approximately baseband waveform. To avoid messy computation, we look at a waveform baseband-limited to 3/2 which is sampled at rate 1 (i.e. sampled at only 1/3 the rate that it should be sampled at). In particular, let $u(t) = \operatorname{sinc}(3t)$.

(a) Sketch $\hat{u}(f)$. Sketch the function $\hat{v}_m(f) = \operatorname{rect}(f-m)$ for each integer m such that $v_m(f) \neq 0$. Note that $\hat{u}(f) = \sum_m \hat{v}_m(f)$.

(b) Sketch the inverse transforms $v_m(t)$ (real and imaginary parts if complex).

(c) Verify directly from the equations that $u(t) = \sum v_m(t)$. [Hint. This is easier if you express the sine part of the sinc function as a sum of complex exponentials.]

(d) Verify the sinc-weighted sinusoid expansion, (4.73). (There are only three nonzero terms in the expansion.)

(e) For the approximation $s(t) = u(0)\,\text{sinc}\,(t)$, find the energy in the difference between $u(t)$ and $s(t)$ and interpret the terms.

Solution: (a) The fourier transforms of $u(t)$

$$\hat{u}(f) = \mathscr{F}[u(t)] = \frac{1}{3}\,\text{rect}\left(\frac{f}{3}\right), \tag{16}$$

as shown in figure 3.

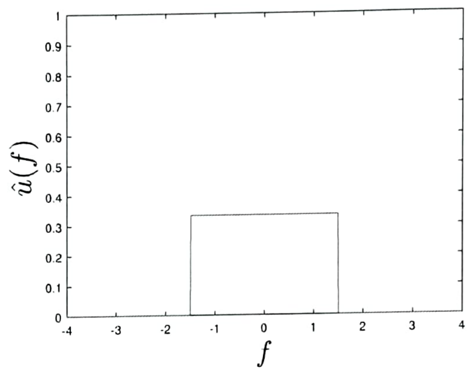


Figure 3: $\hat{u}(f)$

The segment functions are

$$\hat{v}_0(f) = \frac{1}{3}\,\text{rect}\,(f), \tag{17}$$
$$\hat{v}_1(f) = \frac{1}{3}\,\text{rect}\,(f-1), \tag{18}$$
$$\hat{v}_{-1}(f) = \frac{1}{3}\,\text{rect}\,(f+1), \tag{19}$$

as shown in figure 4.

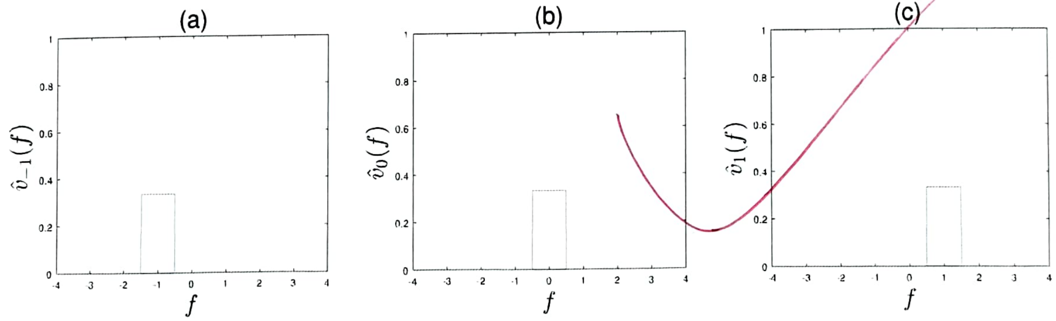


Figure 4: $\hat{v}_m(f)$ for $m = 0, \pm 1$.

(b) The inverse transform of $\hat{v}_0(f)$ is

$$v_0(t) = \mathscr{F}^{-1}[\hat{v}_0(t)] = \frac{1}{3}\operatorname{sinc}(t). \tag{20}$$

as shown in figure 5.

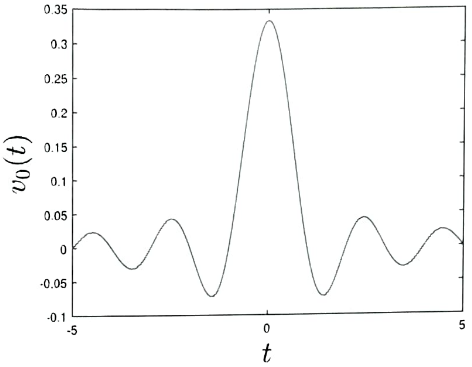


Figure 5: $v_0(t)$.

The inverse transform of $\hat{v}_1(f)$ is

$$v_1(t) = \mathscr{F}^{-1}[\hat{v}_1(t)] = \frac{1}{3}\operatorname{sinc}(t)e^{2\pi i t}, \tag{21}$$

whose real part is

$$\mathrm{Re}\,[v_1(t)] = \frac{1}{3}\operatorname{sinc}(t)\cos(2\pi t), \tag{22}$$

and imaginary part is

$$\mathrm{Im}\,[v_1(t)] = \frac{1}{3}\operatorname{sinc}(t)\sin(2\pi t), \tag{23}$$

as shown in figure 6.

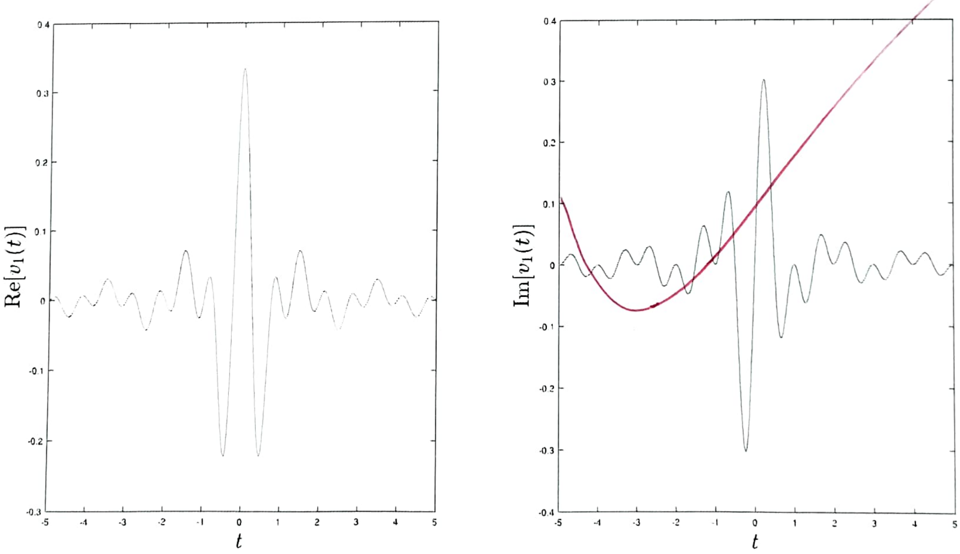


Figure 6: Real part and imaginary part of $v_1(t)$.

The inverse transform of $\hat{v}_{-1}(f)$ is

$$v_{-1}(t) = \mathscr{F}^{-1}[\hat{v}_{-1}(t)] = \frac{1}{3}\operatorname{sinc}(t)e^{-2\pi i t}, \tag{24}$$

whose real part is

$$\operatorname{Re}[v_{-1}(t)] = \frac{1}{3}\operatorname{sinc}(t)\cos(2\pi t), \tag{25}$$

and imaginary part is

$$\operatorname{Im}[v_{-1}(t)] = -\frac{1}{3}\operatorname{sinc}(t)\sin(2\pi t), \tag{26}$$

as shown in figure 7.

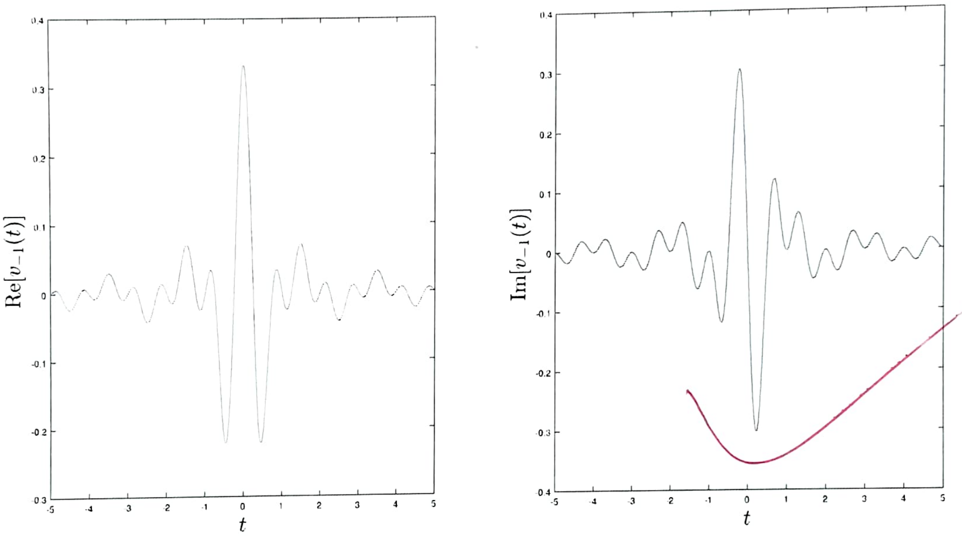


Figure 7: Real part and imaginary part of $v_{-1}(t)$.

(c)

$$\begin{aligned}\sum_m v_m(t) =& v_0(t) + v_1(t) + v_{-1}(t) = \frac{1}{3}\frac{e^{\pi i t} - e^{-\pi i t}}{2\pi i t} + \frac{1}{3}\frac{e^{\pi i t} - e^{-\pi i t}}{2\pi i t}e^{2\pi i t} + \frac{1}{3}\frac{e^{\pi i t} - e^{-\pi i t}}{2\pi i t}e^{-2\pi i t}\\ =& \frac{1}{3}\frac{e^{3\pi i t} - e^{-3\pi i t}}{2\pi i t} = \frac{\sin(3\pi t)}{3\pi t} = \operatorname{sinc}(3t) = u(t)\end{aligned} \tag{27}$$

(d) Sampling period $T = 1$,

$$\begin{aligned}\sum_{m,k} v_m(kT)\operatorname{sinc}\left(\frac{t}{T} - k\right)e^{2\pi i m t/T} =& \sum_{m=0,\pm 1} v_m(0)\operatorname{sinc}(t)e^{2\pi i m t} = \frac{1}{3}\sin(t) + \frac{1}{3}\operatorname{sinc}(t)e^{2\pi i t} + \frac{1}{3}\operatorname{sinc}(t)e^{-2\pi i t}\\ =& \frac{1}{3}\frac{e^{\pi i t} - e^{-\pi i t}}{2\pi i t} + \frac{1}{3}\frac{e^{\pi i t} - e^{-\pi i t}}{2\pi i t}e^{2\pi i t} + \frac{1}{3}\frac{e^{\pi i t} - e^{-\pi i t}}{2\pi i t}e^{-2\pi i t}\\ =& \frac{1}{3}\frac{e^{3\pi i t} - e^{-3\pi i t}}{2\pi i t} = \frac{\sin(3\pi t)}{3\pi t} = \operatorname{sinc}(3t) = u(t)\end{aligned}$$

(e) The approximation function

$$s(t) = u(0)\operatorname{sinc}(t) = \operatorname{sinc}(t). \tag{28}$$

Using Parseval's Theorem, the energy between $u(t)$ and $s(t)$ is

$$\int_{-\infty}^{+\infty} |u(t) - s(t)|^2 \, \mathrm{d}t = \int_{-\infty}^{+\infty} |\hat{u}(f) - \hat{s}(f)| \; \mathrm{d}f = \int_{-\infty}^{+\infty} \left| \frac{1}{3} \operatorname{rect}\left(\frac{f}{3}\right) - \operatorname{rect}(f) \right|^2 \mathrm{d}f = \frac{2}{3}.$$

which shows that $s(t)$ is not a very good approximation of $u(t)$, i.e., $s(t)$ and $u(t)$ are not $\mathcal{L}_2$ equivalent. □

25