NUS

National University of Singapore

Name : LUO ZIJIAN

Matric.No： A0224725H

MUSNET： E0572844

Subject： Information Theory

Assignment: Homework Four

Date: Sep 5^{th}

Prof: Marco Tomamichel

EXERCISE 4.1

(a) For the entropy of X, $H(x) = \sum_{x\in X} P(x=i) \log \frac{1}{P(x=i)}$

$$= \sum_{x\in X} 2^{-i} \log 2^{i}$$

$$= \sum_{x\in X} i \cdot 2^{-i}$$

$$H(x) = 1\times 2^{-1} + 2\times 2^{-2} + 3\times 2^{-3} + \cdots\ i\cdot 2^{-i} \quad \cdots\ ①$$

$$\frac{1}{2} H(x) = \quad + 1\times 2^{-2} + 2\times 2^{-3} + \cdots (i-1)\cdot 2^{-i} + i\cdot 2^{-i-1} \cdots ②$$

① − ②, we get $\frac{1}{2}H(x) = 1 - (\frac{i}{2}+1)(\frac{1}{2})^{i}$

$$H(x) = 2 - (i+2)(\tfrac{1}{2})^{i}$$

$$= 2 - \frac{i+2}{2^{i}}$$

$$\cancel{\quad} = 2$$

(b). For Huffman coding, we set $P(x=1) = 2^{-1} \Rightarrow 0$

$P(x=2) = 2^{-2} \Rightarrow 10$

$\vdots$

$P(x=i-1) = 2^{-(i-1)} \Rightarrow 111\cdots0$

$P(x=i) = 2^{-i} \Rightarrow 111\cdots1$

Expected length = $1\times 2^{-1} + 2\times 2^{-2} + \cdots + (i-1)\times 2^{-(i-1)} + (i-1)\times 2^{-i}$

We know the entropy of x [from part (a)]

$$H(x) = 1\times 2^{-1} + 2\times 2^{-2} + \cdots + (i-1)\times 2^{-(i-1)} + i\times 2^{-i}$$

Expected length − H(X) = -2^{-i}

so Expected length ≤ H(x) ... ①

And at the same time Expected length ≥ H(x) ... ②

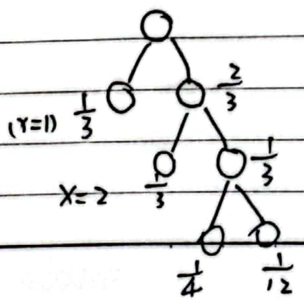
From ①,②, so Expected length = H(X)

In a word, it is indeed optimal code.

Exercise 4.2

a). Through the order from $(\frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{12})$

$Pr(X=1) = \frac{1}{3}$, $Pr(X=2) = \frac{1}{3}$, $Pr(X=3) = \frac{1}{4}$, $Pr(X=4) = \frac{1}{12}$



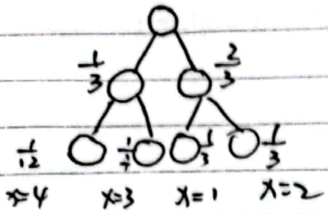
$x=1 \Rightarrow 0$

$x=2 \Rightarrow 10$

$x=3 \Rightarrow 110$

$x=4 \Rightarrow 111$

(b) However, it exists a different optimal set



$x=1 \Rightarrow 10$
$x=2 \Rightarrow 11$
$x=3 \Rightarrow 01$
$x=4 \Rightarrow 00$

$H(x) = 1.855$ bit

Therefore (1, 2, 3, 3) and (2, 2, 2, 2) all exists.

Next, we prove why these two codeword length assignments are both optimal.

For (1, 2, 3, 3), we get expected length $1\times\frac{1}{3} + 2\times\frac{1}{3} + 3\times\frac{1}{4} + 3\times\frac{1}{12} = 2$ bits

$1.855 \text{ bits} < 2 \text{ bits} < 1.855+1 \text{ bits}$

Therefore, it is optimal

For (2, 2, 2, 2), we get expected length $2\times\frac{1}{3} + 2\times\frac{1}{3} + 2\times\frac{1}{4} + 2\times\frac{1}{12} = 2$ bits.

with same calculation, we also can adust this assignment lengths set is also optimal.

c) No, there are no any optimal codes with codeword lengths can exceed the shannon code length $\lceil \log \frac{1}{p(x)} \rceil$

For shannon code: $\frac{1}{3}\times\lceil\log_2 3\rceil + \frac{1}{3}\times\lceil\log_2 3\rceil + \frac{1}{4}\times\lceil\log_2 4\rceil + \frac{1}{12}\times\lceil\log 12\rceil$

$= \frac{2}{3} + \frac{2}{3} + \frac{1}{2} + \frac{1}{3}$

$= 2.166$ bit

Obviously, 2.166 bits > 2 bits

Therefore, there are no any optimal codes with codeword lengths can exceed the shannon code length $\lceil \log \frac{1}{p(x)} \rceil$

Exercise 4.3

a) Algorithm: we choose most frequent laters with codewords of length which costs less time and less length

For example, "." $\Rightarrow$ 0, "-" $\Rightarrow$ 1

time(2) $\Rightarrow$ 0

time(4) $\Rightarrow$ 00, 1

time(6) $\Rightarrow$ 000, 01, 10

time(8) $\Rightarrow$ 0000, 001, 010, 100, 11

time(10) $\Rightarrow$ 00000, 0001, 0010, 0100, 1000, 011, 101, 110

time(12) $\Rightarrow$ 0011, 0101, 0110, 1001, 1010, 1110, 111

Therefore, the order is 0, 1, 00, 01, 10, 000, 11, 001, 010, 100, 0000, 011, 110, 0001, 0010, 0100, 1000, 00000, 111, 0011, 0101, 0110, 1001, 1010, 1110

b). After calculating the result, we get average time 9.7140

c). After calculating the result, we get average length: 4.484 bits

$H(x) = 3.90$ bit

$H(x)$ < expected average length < $H(x)+1$

Therefore, it is optimal

d). Prefix code:

Exercise 4.3 (2)

```
%%%Exercise 4.3(2) part
pp=[8.4,1.5,2.2,4.2,11,2.2,2.0,6,7.4,0.1,1.3,4,2.4,
6.7,7.4,1.9,0.1,7.5,6.2,9.2,2.7,0.9,2.5,0.1,2,0.1];
A='a':'z';
AA=mat2cell(A,1,ones(26,1));
AAA=[];
%sort probabilities and record the index
[pp0,pindex]=sort(pp,'descend');
%generate codewords
code={'0','1','00','01','10','000','11','001','010'
,'100','0000','011','101','110','0001','0010','0100
','1000','00000','111','0011','0101','0110','1001',
'1010','1110'};
%combine symbols and codewords
for i=1:26
   AA(2,pindex(i))=cellstr([cell2mat(code(i)),'_']);
end
%expected length
Sum=0;
for i=1:26
    Sum=Sum+pp(i)*length(cell2mat(AA(2,i)));
end
Sum=Sum/100;
%expected time
time=0;
for i=1:26
    t0=2*length(find(cell2mat(AA(2,i))=='0'));
    t1=4*length(find(cell2mat(AA(2,i))=='1'));
    time=time+pp(i)/100*(t0+t1);
end
time=time+3
```

Exercise 4.3 (3)

```
pp=[8.4,1.5,2.2,4.2,11,2.2,2.0,6,7.4,0.1,1.3,4,2.4,
6.7,7.4,1.9,0.1,7.5,6.2,9.2,2.7,0.9,2.5,0.1,2,0.1];
Sum=0;
for i=1:26
    p=pp(i)/100;
    Sum=Sum+p*log(1/p);
end
Sum
```

Exercise 4.4 $H(X,Y) = \lim_{n\to\infty} \frac{1}{n} H(X^1, X^2 \dots X^n, Y^1, Y^2 \dots Y^n)$

$H(X) = \lim_{n\to\infty} \frac{1}{n} H(X^1, X^2, \dots X^n)$

~~For (1)~~

$H(Y) = \lim_{n\to\infty} \frac{1}{n} H(Y^1, Y^2, \dots Y^n)$

$H(X|Y) = H(X,Y) - H(Y)$ $\quad R(Y|X) = H(X,Y) - H(X)$

And, we also know $\lim_{n\to\infty} P\{(\hat{X}^n, \hat{Y}^n) \neq (X^n, Y^n)\} = 0$

From the conclusion $[R X_1 \geq H(X)]$ is achievable. ... ①

For (1)

$$H(X) - H(X|Y) = \sum_{x\in X} P_x \log \frac{1}{P_x} - \sum_{x\in X, Y\in Y} P_{X|Y} \log \frac{1}{P_{X|Y}}$$

$$= \sum_{X\in X, Y\in Y} P_Y \cdot P_{X|Y} \log \frac{1}{P_{X|Y} P_Y} - \sum_{X\in X, Y\in Y} P_{X|Y} \log \frac{1}{P_{X|Y}}$$

$$= \sum_{X\in X, Y\in Y} P_{X|Y} \log \frac{P_{X,Y}}{P_{X|Y}} \boxed{\cancel{P_Y}}.$$

~~The~~ Therefore $H(X) - H(X|Y) \geq 0$ ... ②

Combined with ① and ②, we got $R_1 \geq H(X|Y)$

For (2)

This proof is similar, so, we also can get $R_2 \geq H(Y|X)$

For (3) $R_1 + R_2 \geq H(X) + H(Y)$ ..... ③

$$H(X) + H(Y) - H(X,Y) = \sum_{X\in X} P_X \log \frac{1}{P_X} + \sum_{Y\in Y} P_Y \log \frac{1}{P_Y} - \sum_{X\in X}\sum_{Y\in Y} P_{X,Y} \log \frac{1}{P_{X,Y}}$$

$$= \sum_{X\in X}\sum_{Y\in Y} \log \frac{P_{X,Y}}{P_X \cdot P_Y}.$$

$$\geq 0$$

so, $H(X) + H(Y) \geq H(X,Y)$ ... ④

Combined with ③ and ④, we can get $R_1 + R_2 \geq H(X,Y)$