Fundamentals Of Information Science **2022 Spring**

Homework5

学生: 华园（202000120027）） *时间: 2022.3.23*

Problem 1.

Lempel-Ziv
Give the LZ77 parsing and encoding of 000000110101000001101 01. Let the window size W be 8.

Solution.

$$\begin{array}{rl} & 0，00000，1，1，01，010，0000，1，10，101 \\ \text{LZ77:} & (0,0)(1,1,5)(0,1)(1,1,1)(1,3,2)(1,2,3)(1,1,4)(1,6,1)(1,7,2)(1,2,3) \end{array}$$

Problem 2.

Consider the LZW compression and decompression algorithms. Assume that the scheme has an initial table with codewords 0 through 255 corresponding to the 8-bit ASCII characters; character "a" is 97 and "b" is 98 . The receiver gets the following sequence of codewords, each of which is 10 bits long:

97 97 98 98 257 256

(a)What was the original message sent by the sender?
(b) By how many bits is the compressed message shorter than the original message(each character in the original message is 8 bits long)?
(c)What is the first string of length 3 added to the compression table?(If there's no such string,your answer should be "None".)

Solution.

$$\begin{array}{cccccc} 97 & 97 & 98 & 98 & 257 & 256 \\ | & | & | & | & | & | \\ 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 'a' & 'a' & 'b' & 'b' & 'ab' & 'aa' \end{array}$$

256	aa
257	ab
258	bb
259	ba
260	aba
261	none

(a)the original message sent by the sender is 'aabbabaa'

$$\begin{aligned} original\ message: &= 8*8 = 64\ (bits) \\ compressed\ message: &= 6*10 = 60\ (bits) \\ 64 - 60 &= 4\ (bits) \end{aligned}$$

(b)Therefore, a total of 4 bits is shortened.
(c)'aba',which is on line 260 of the table

Problem 3.

Let $X = HHTHT$ be a sequence generated from a biased coin with an unknown probability, you are asked to show:
(1) how to generate random bits from X if applying the von Veumann's algorithm;
(2) how to generate random bits from X if applying the Elias's algorithm.

Since I don't fully understand the meaning of the question, I give two ideas:
The first idea:

Solution.

(1)applying the von Veumann's algorithm:
Let HT represent 0 and TH represent 1:

HHTHT can be regarded as x00(x stands for abandonment)
Therefore, 2 bits are generated

HHTHT can also be regarded as x1x(x stands for abandonment)
Therefore, 1 bit are generated

(2)applying the Elias's algorithm. Let HTHT represent 01

HHTHT can be regarded as x01(x stands for abandonment)
Therefore, 2 bits are generated

The Second idea:

Solution.

(1)applying the von Veumann's algorithm:

Let HHTHT represent 0 and HHTTH represent 1:

We can therefore generate 1 bit crandom numbers.

(2)applying the Elias's algorithm.
Sequence with equal probability to X:

HHHTT
HHTHT
HTHHT
THHHT
HHTTH
HTHTH
THHTH
HTTHH
THTHH
TTHHH

We can use the first eight sequences to represent (000,001,.......111) respectively,which generate 3bits.We can use the last two sequences to represent (0, 1) respectively,which generate 1 bit.

Problem 4.

Z-channel. The Z-channel has binary input and output alphabets and transition probabilities $p(y \mid x)$ given by the following matrix:

$$Q = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix} \quad x, y \in \{0, 1\}$$

Find the capacity of the Z-channel and the maximizing input probability distribution.

Solution.

We assume that $P(x = 0) = P, P(x = 1) = 1 - P$:

$$P(y = 0) = \frac{1 + P}{2}$$
$$P(y = 1) = \frac{1 - P}{2}$$

Then there:

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y \mid X) \\ &= H(Y) - (1 - P) \\ &= -\left\{\frac{1}{2}(1 + P)\log_2\left[\frac{1}{2}(1 + P)\right] + \frac{1}{2}(1 - P)\log_2\left[\frac{1}{2}(1 - P)\right] + (1 - p)\right\} \\ I'(X;Y) &= -\left\{\frac{1}{2}\log_2\left(\frac{1 + P}{1 - P}\right) - 1\right\} \end{aligned}$$

Let I'(X;Y)=0,we can get when P=0.6,I'(X,Y)=0:

when p<0.6	I'(X,Y)>0
when p>0.6	I'(X,Y)<0

$$I(X;Y)_{max} = I(X;Y)|_{P=0.6} = -\left(\frac{4}{5}\log_2\frac{4}{5} + \frac{1}{5}\log_2\frac{1}{5} + \frac{2}{5}\right) \approx 0.322$$

the capacity of the Z-channel is 0.322，the maximizing input probability distribution is $X = \left\{ \begin{array}{ll} 0 & \mathrm{P} = 0.6 \\ 1 & \mathrm{P} = 0.4 \end{array} \right.$