German University in Cairo - GUC
Faculty of Information Engineering & Technology - IET
Department of Communication Engineering
Winter Semester 2013/2014

Students Name: **Students ID:**

COMM901 – Source Coding and Compression

Quiz 1

		possible points	achieved points
Question 1	(a)	3	
	(b)	2	
	(c)	3	
	total	**8**	
Question 2	(a)	4	
	(b)	5	
	(c)	3	
	total	**12**	
Question 3	(a)	4	
	(b)	3	
	(c)	1	
	total	**8**	
Total		**25 (+3 bonus)**	
Percentage			

Question 1: Kraft inequality and prefix codes [8 points]

Given is a stationary discrete random process $\boldsymbol{S} = \{S_n\}$ with independent and identically distributed random variables S_n. The alphabet $\mathcal{A}_S$ for the random variables S_n includes six letters and is given by $\mathcal{A}_S = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \mathrm{e}, \mathrm{f}\}$.

(a) The following table shows three sets of codeword lengths for the given alphabet $\mathcal{A}_S$. The codeword length assignments are denoted as "set A", "set B" and "set C". An assignment $\ell(a_i) = M$ describes that, for a selected codeword set, a codeword of M bits is assigned to the alphabet letter a_i.

a_i	assignment of codeword lengths $\ell(a_i)$		
	set A	set B	set C
a	2	2	1
b	2	2	3
c	2	2	3
d	4	3	3
e	4	3	4
f	4	4	4

Determine for each of the given sets, using the Kraft inequality, whether it is possible to construct a uniquely decodable code with the corresponding codeword lengths. [3 points]

Set A: $$\sum_{\forall a_i} 2^{-\ell(a_i)} = 3 \cdot 2^{-2} + 3 \cdot 2^{-4} = \frac{3}{4} + \frac{3}{16} = \frac{12+3}{16} = \frac{15}{16} \leq 1$$

$\Longrightarrow$ A uniquely decodable code can be constructed for set A. [1]

Set B: $$\sum_{\forall a_i} 2^{-\ell(a_i)} = 3 \cdot 2^{-2} + 2 \cdot 2^{-3} + 2^{-4} = \frac{3}{4} + \frac{2}{8} + \frac{1}{16} = \frac{12+4+1}{16} = \frac{17}{16} > 1$$

$\Longrightarrow$ A uniquely decodable code cannot be constructed for set B. [1]

Set C: $$\sum_{\forall a_i} 2^{-\ell(a_i)} = 2^{-1} + 3 \cdot 2^{-3} + 2 \cdot 2^{-4} = \frac{1}{2} + \frac{3}{8} + \frac{2}{16} = \frac{4+3+1}{8} = \frac{8}{8} \leq 1$$

$\Longrightarrow$ A uniquely decodable code can be constructed for set C. [1]

(b) Develop a prefix code for the set of codeword lengths given in the table below. The codeword that is assigned to a letter a_i shall have $\ell(a_i)$ bits. Write the codewords of the developed code directly into the table. [2 points]

a_i	$\ell(a_i)$	codewords
a	1	0
b	3	100
c	3	101
d	3	110
e	4	1110
f	4	1111

[2]

(c) Consider the prefix code developed in (b). Is it possible to find a pmf $p(a_i)$ for which the developed code yields an average codeword length $\bar{\ell}$ that is equal to the entropy (S_n) ? Briefly explain your answer. If such a pmf exists, write down a table with the corresponding probability masses. [3 points]

<u>Note:</u> The codeword lengths for the prefix code developed in (b) are equal to the codeword lengths for "set C", which has been investigated in (a).

The average codeword length $\bar{\ell}$ for a code is equal to the entropy $H(S_n)$ if and only if the lengths of the codewords that are assigned to the alphabet letters obey the condition $\ell(a_i) = -\log_2 p(a_i)$. This is only possible if all probability masses $p(a_i)$ determined in this way add up to 1, or in other words, if the Kraft inequality is fulfilled with equality. [1]

It has been shown in (a), that the set of codeword lengths fulfills the Kraft inequality with equality. Hence, it is possible to find a pmf yielding $\bar{\ell} = H(S_n)$. To achieve that the probability masses have to be set according to $p(a_i) = 2^{-\ell(a_i)}$. [1]

The corresponding probability masses are listed in the following table.

a_i	$p(a_i)$
a	1/2 = 0.5
b	1/8 = 0.125
c	1/8 = 0.125
d	1/8 = 0.125
e	1/16 = 0.0625
f	1/16 = 0.0625

[1]

Question 2: Huffman codes [12 points]

The given source $\boldsymbol{S} = \{S_n\}$ represents a Bernoulli process (the random variables S_n are binary, independent and identically distributed) and has the symbol alphabet $\mathcal{A}_S = \{\mathrm{x}, \mathrm{o}\}$. The probability that a random variable S_n takes the value "x" is $P(S_n = \text{"x"}) = 0.8$.

Two codes shall be investigated for coding messages of the given source. Code A assigns a codeword to two successive symbols of a message, whereas code B assigns codewords to variable-length sequences of symbols.

(a) Determine, for both codes, the probability masses for the given symbol sequences $\boldsymbol{s}_i$. Assign codewords to the given symbol sequences $\boldsymbol{s}_i$ so that the resulting code is uniquely decodable (for arbitrary long messages) and yields the smallest possible average codeword length per symbol. Write the determined probability masses and codewords directly into the tables. [4 points]

code A

$\boldsymbol{s}_i$	$p(\boldsymbol{s}_i)$	codeword
oo	0.04	000
ox	0.16	001
xo	0.16	01
xx	0.64	1

[2]

code B

$\boldsymbol{s}_i$	$p(\boldsymbol{s}_i)$	codeword
o	0.200	00
xo	0.160	010
xxo	0.128	011
xxx	0.512	1

[2]

Since the random variables S_n are independent, the probability masses $p(\boldsymbol{s}_i)$ for the symbol sequences $\boldsymbol{s}_i$ are given by the products of the corresponding letter probabilities. A uniquely decodable code that yields the smallest possible average codeword length is obtained, if the codewords are assigned using the Huffman algorithm.

(b) Which of the two developed codes, code A and code B, has a higher coding efficiency for the given source? [5 points]

The average codeword length per symbol for code A is given by:

$$\overline{\ell_A} = \frac{1}{2}\sum_{i=0}^{3} p(\boldsymbol{s}_i)\cdot\ell(\boldsymbol{s}_i) \quad [1]$$
$$= \frac{1}{2}(0.04\cdot 3 + 0.16\cdot 3 + 0.16\cdot 2 + 0.64\cdot 1) = \frac{1.56}{2} = 0.78 \quad [1]$$

The average codeword length per symbol for code B is given by:

$$\overline{\ell_B} = \frac{\sum_{i=0}^{3} p(\boldsymbol{s}_i)\cdot\ell(\boldsymbol{s}_i)}{\sum_{i=0}^{3} p(\boldsymbol{s}_i)\cdot n(\boldsymbol{s}_i)} \quad [1]$$
$$= \frac{0.2\cdot 2 + 0.16\cdot 3 + 0.128\cdot 3 + 0.512\cdot 1}{0.2\cdot 1 + 0.16\cdot 2 + 0.128\cdot 3 + 0.512\cdot 3} = \frac{1.776}{2.44} \approx 0.727869 \quad [1]$$

Since $\overline{\ell_B} < \overline{\ell_A}$, the code B has a higher coding efficiency for the given source. [1]

(c) The more efficient of the codes A and B is used for coding very long messages of the given source. What is the maximum percentage of bit rate that can be saved on average if the used code is replaced by more efficient lossless coding techniques? [3 points]

The minimum required average codeword length per symbol is given by the entropy rate of the source. Since the given source is a stationary source with independent random variables, the entropy rate is equal to the marginal entropy:

$$\begin{aligned}\bar{H}(\boldsymbol{S}) &= H(S_n)\\ &= -p(\text{"x"})\cdot\log_2 p(\text{"x"}) - p(\text{"o"})\cdot\log_2 p(\text{"o"}) \quad [1]\\ &= -0.8\cdot\log_2 0.8 - 0.2\cdot\log_2 0.2\\ &\approx 0.721928 \quad [1]\end{aligned}$$

Since code B is the more efficient code, the maximum percentage of the bit rate that can be saved on average is given by

$$\rho = \frac{\overline{\ell_B} - \bar{H}(\boldsymbol{S})}{\overline{\ell_B}} = 1 - \frac{\bar{H}(\boldsymbol{S})}{\overline{\ell_B}} \approx 1 - \frac{0.721928}{0.727869} \approx 0.008162$$

Compared to the more efficient code B, at maximum 0.82 % of the bit rate can be saved on average by more efficient lossless coding techniques. [1]

Question 3: Entropy and Entropy rate [8 points]

The instantaneous entropy rate $\overline{H}_{inst}(\boldsymbol{S}, N)$ of a stationary random process $\boldsymbol{S} = \{S_n\}$ is defined according to

$$\overline{H}_{inst}(\boldsymbol{S}, N) = \frac{1}{N} \cdot H_N(S_k, S_{k+1}, \cdots, S_{k+N-1})$$

where $H_N(S_k, S_{k+1}, \cdots, S_{k+N-1})$ denotes the block entropy of N successive random variables. The instantaneous entropy rate $\overline{H}_{inst}(\boldsymbol{S}, N)$ specifies a lower bound for the average codeword length for the coding of messages of exactly N symbols. It depends on the parameter N.

(a) Proof that, for stationary Markov processes, the instantaneous entropy rate can also be written as

$$\overline{H}_{inst}(\boldsymbol{S}, N) = \frac{1}{N} \cdot H(S_n) + \frac{N-1}{N} \cdot H(S_n|S_{n-1})$$

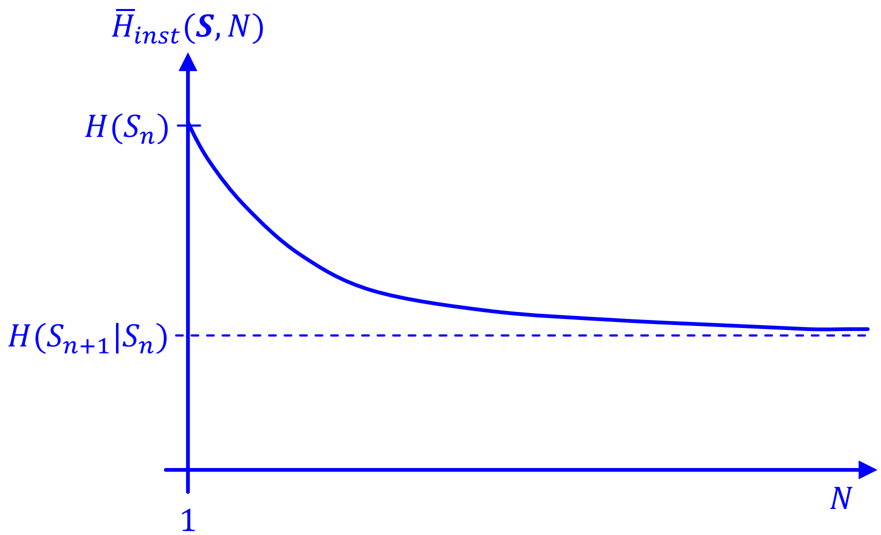
where $H(S_n)$ represents the marginal entropy and $H(S_n|S_{n-1})$ represents the conditional entropy given the last symbol. [4 points]

Given the definition above, the instantaneous entropy rate for stationary Markov processes can be written as:

$$\begin{aligned}
\overline{H}_{inst}(\boldsymbol{S}, N) &= \frac{1}{N} \cdot H_N(S_k, S_{k+1}, \cdots, S_{k+N-1}) \\
&= \frac{1}{N} \cdot E\{-\log_2 p(S_k, S_{k+1}, \cdots, S_{k+N-1})\} \quad [1] \\
&= \frac{1}{N} \cdot E\{-\log_2 (p(S_k) \cdot p(S_{k+1}|S_k) \cdot p(S_{k+2}|S_{k+1}) \ldots p(S_{k+N-1}|S_{k+N-2}))\} \quad [1] \\
&= \frac{1}{N} \cdot E\{-\log_2 (p(S_n) \cdot p(S_{n+1}|S_n)^{N-1})\} \quad [1] \\
&= \frac{1}{N} \cdot E\{-\log_2 p(S_n) - (N-1) \cdot \log_2 p(S_{n+1}|S_n)\} \\
&= \frac{1}{N} \cdot E\{-\log_2 p(S_n)\} - \frac{N-1}{N} \cdot E\{-\log_2 p(S_{n+1}|S_n)\} \\
&= \frac{1}{N} \cdot H(S_n) + \frac{N-1}{N} \cdot H(S_n|S_{n-1}) \quad [1]
\end{aligned}$$

Here, we used the fact that, for Markov processes, the joint pmf $p(S_k, S_{k+1}, \cdots, S_{k+N-1})$ can be written as the product $p(S_k) \cdot p(S_{k+1}|S_k) \cdot p(S_{k+2}|S_{k+1}) \cdot \ldots \cdot p(S_{k+N-1}|S_{k+N-2})$. And we used the fact that, for stationary processes, all conditional probabilities $p(S_{k+i+1}|S_{k+i})$ are the same.

(b) Sketch the instantaneous entropy rate $\overline{H}_{inst}(\boldsymbol{S}, N)$ for a general stationary Markov process as a function of N. Label the value for $N = 1$ and the limit for $N \to \infty$. [3 points]



[3]

(c) For which stationary processes is the instantaneous entropy rate $\overline{H}_{inst}(\boldsymbol{S}, N)$ independent of the parameter N? An explanation is not required. [1 point]

The instantaneous entropy rate is independent of the parameter N if the random variables of the stationary process are independent, i.e., if the source represents an iid process. [1]