Lund University
Electrical and Information Technology

Exam in
Information Theory, EIT080

August 26, 2011, kl 8-13

- Write your name on each paper.
- Start a new solution on a new sheet of paper. Use only one side of the paper.
- Solutions should clearly show the line of reasoning.
- Aid:
  - Course book.
  - Handed out material.
  - Lecture slides and articles from the home page.
  - Calculator.

Good luck!

Problem 1

The joint probability for the random variables X and Y is given by the following table

$P(X,Y)$		$Y=0$	$Y=1$
X	0	$1/8$	$3/8$
	1	$1/4$	0
	2	$1/8$	$1/8$

Calculate

(a) $H(X)$ and $H(Y)$

(b) $H(X|Y)$ and $H(Y|X)$

(c) $H(X,Y)$ and $I(X;Y)$

(3+4+3=10p)

Problem 2

Consider the discrete memoryless channel in Figure 2.1

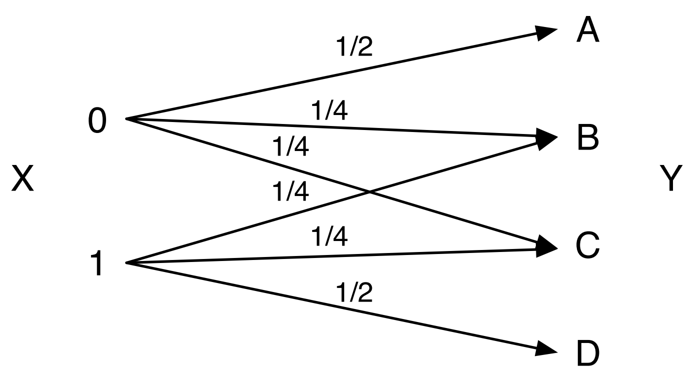

Figure 2.1: A discrete memoryless channel.

(a) What is the channel capacity and for what distribution on X is it reached?

(b) Assume now that $P(X = 0) = 1/6$ and $P(X = 1) = 5/6$ and that the source is memoryless. Find an optimal code to compress a sequence $\boldsymbol{Y}$. What is the average codeword length?

(5+5=10p)

Problem 3

In Figure 3.1 a graph for a Markov source is shown. Derive the stationary distribution for the source and the entropy rate $H_\infty(X)$.

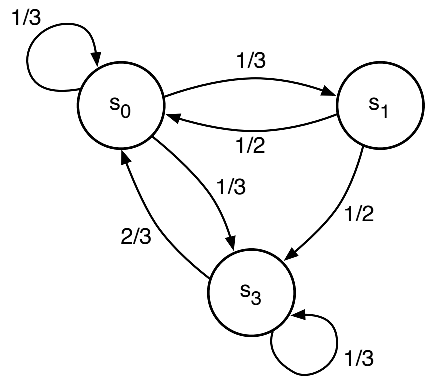

Figure 3.1: A graph for a Markov process.

(10p)

Problem 4

Consider the sequence

`the_rain_in_spain_falls_mainly_in_the_plain`

where '_' is used to denote a space. Use an LZSS algorithm to encode the sequence. The LZSS is an LZ77 algorithm with the improvement that the uncoded letter is only sent if it is not found in the search buffer. A binary prefix is used to separate the two different kind of codewords:

- $(0, j, \ell)$ denotes that the first ℓ symbols of the look ahead buffer is found on position j in the search buffer.
- $(1, x)$ denotes that the first letter x in the look ahead buffer is not found in the search buffer.

Assume that the search buffer is of length $S = 16$ and the look ahead buffer is $B = 8$ characters long. That means that j is encoded with four bits and ℓ with three bits. Assume that all characters are encoded according to the ASCII table, i.e. with eight bits each. What is the compression rate for this sequence?

(10p)

Problem 5

Consider a band-limited Gaussian channel with bandwidth W. Then the transmitted signal $X(t)$ is distorted by the additive noise $Z(t) \in \mathcal{N}(0, \sqrt{N_0/2})$ and followed by low-pass filter with bandwidth W to form

$$Y(t) = \big(X(t) + Z(t)\big) * h_{LP}(t)$$

The signal is sampled with frequency $F_s = 2W$ samples/s. We know from the theory that the capacity for this channel is

$$C_t = W \log\Big(1 + \frac{P}{N_0 W}\Big) \quad \text{[bit/s]}$$

where $P = E_b R_t$ and R_t the information rate. From this result we can also find, by letting $W \to \infty$, the fundamental Shannon limit, stating that reliable communication is possible if and only if the signal to noise ratio fulfill

$$\frac{E_b}{N_0} > \ln 2 = -1.59 \text{ dB}$$

To get even close to this limit the system must use coding. We can assume codewords of length N samples corresponding to K information bits, yielding a code rate of $R = K/N$.

However, the capacity result above requires that the coding rate tend to zero. Here we will consider the requirements on the signal to noise ratio E_b/N_0 for arbitrary code rates.

(a) Show that the information rate (information bits per second) can be written[1]

$$R_t = 2WR \quad \text{[bit/s]}$$

and that

$$\frac{P}{N_o W} = 2\frac{E_b}{N_0} R$$

(b) Show that for code rate R it is possible to achieve reliable communication if and only if

$$\frac{E_b}{N_0} > \frac{2^{2R} - 1}{2R}$$

Derive the Shannon limit when $R = 1/2$ and when $R \to 0$?

(4+6=10p)

[1] Hint: Let T be the time for a code symbol and derive N.