University of Amsterdam

Faculty of Science

Exam

Information Theory

Master of Logic (5314INTH6Y)

Final exam
Date: Friday, 18 December, 2015
Time: 9:00-12:00

Number of pages: 5 (including front page)
Number of questions: 7
Maximum number of points to earn: 9
At each question is indicated how many points it is worth.

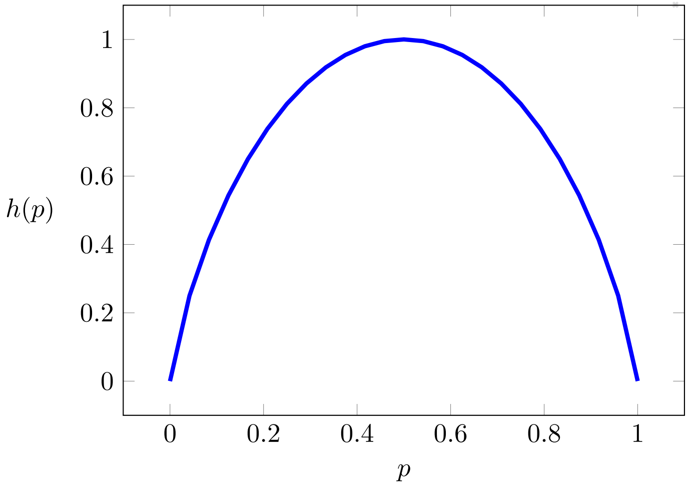
BEFORE YOU START

- Please **wait** until you are instructed to open the booklet.
- Check if your version of the exam is complete.
- Write down **your name, student ID number**, and if applicable the **version number** on **each sheet** that you hand in. Also **number the pages**.
- Your **mobile phone** has to be switched off and in the coat or bag. Your **coat and bag** must be under your table.
- **Tools allowed**: the two course books [CT, MacKay] or printouts of them, printout of script [CF], notes, scratch paper.

PRACTICAL MATTERS

- The first 30 minutes and the last 15 minutes you are not allowed to leave the room, not even to visit the toilet.
- You are obliged to identify yourself at the request of the examiner (or his representative) with a proof of your enrollment or a valid ID.
- During the examination it is not permitted to visit the toilet, unless the proctor gives permission to do so.
- 15 minutes before the end, you will be warned that the time to hand in is approaching.
- If applicable, please fill out the evaluation form at the end of the exam.

Good luck!

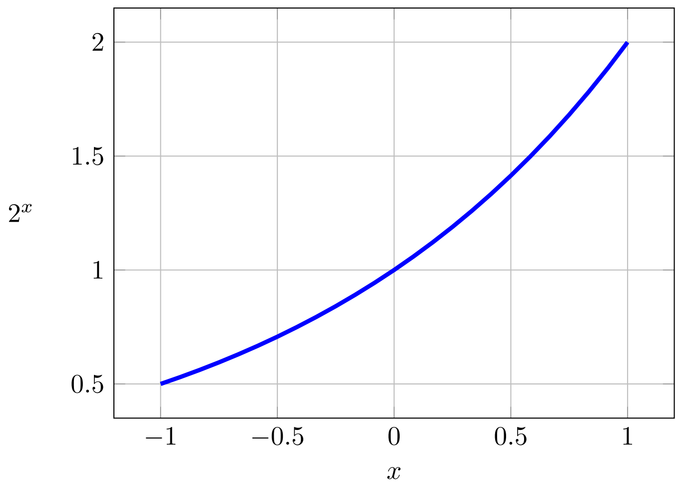


$$h(1/3) = h(2/3) \approx 0.92$$
$$h(1/4) = h(3/4) \approx 0.81$$
$$\frac{1}{4}h(1/3) = \frac{1}{4}h(2/3) \approx 0.23$$
$$\frac{2}{4}h(1/3) = \frac{2}{4}h(2/3) \approx 0.46$$
$$\frac{3}{4}h(1/3) = \frac{3}{4}h(2/3) \approx 0.69$$
$$\frac{1}{3}h(1/4) = \frac{1}{3}h(3/4) \approx 0.27$$
$$\frac{2}{3}h(1/4) = \frac{2}{3}h(3/4) \approx 0.54$$

Figuur 1: Binary entropy function $h(p)$ and some useful approximations.

x	1	2	3	4	5	6	7	8	9	10
$\log_2(x)$	0	1	1.58	2	2.32	2.58	2.81	3	3.17	3.32

Tabel 1: Some useful approximations to the binary logarithm.



x	2^x	2^{-x}
0	1.00	1.00
0.1	1.07	0.93
0.2	1.15	0.87
0.3	1.23	0.81
0.4	1.32	0.76
0.5	1.41	0.71
0.6	1.52	0.66
0.7	1.62	0.62
0.8	1.74	0.57
0.9	1.87	0.54
1	2.00	0.50

Figuur 2: Some useful approximations to some powers of 2.

1. [1.5 points] A random variable X follows the distribution given by the following table:

x	A	B	C	D	E
$P_X(x)$	.3	.3	.2	.1	.1

   (a) Compute $H(X)$. **Hint:** Use the approximations on the previous page.

   (b) Construct a binary Huffman code for the variable.

   (c) Compute the expected codeword length for your code.

   (d) Encode the string `CABBED` according to your code.

   (e) Decode the string `101100110111` according to your code. (Note that there might be a few undecoded bits left over at the end of the string which do not add up to a full codeword. If so, simply ignore these remaining bits.)

   (f) You generate a file

$$X_1, X_2, X_3, \ldots, X_{1200}$$

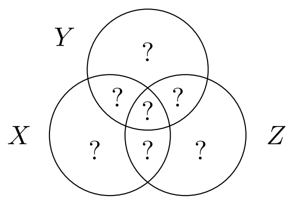
   by drawing 1200 samples i.i.d. from P_X and encode this file according to your code. What is the expected number of `1`s in the output?

2. [1.5 points] Let X, Y, Z be binary random variables with the following joint distribution:

x	y	z	$P_{XYZ}(x, y, z)$
0	0	0	1/4
1	1	0	1/4
0	0	1	1/4
1	0	1	1/4
otherwise			0

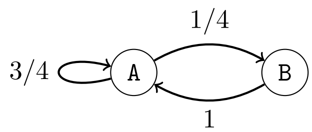
   (a) Give the joint distributions P_{XY}, P_{XZ}, P_{YZ} and the marginals P_X, P_Y, P_Z.

   (b) Draw an entropy diagram as outlined below, and compute and fill in the correct numbers for the question marks:



3. [1 point] Let P_X be a probability distribution with an entropy of $H(X) = 10$ bits, and let K be an optimal source code for this distribution (e.g., a Huffman code). Find an upper bound (strictly below 3/4) on the probability that a sample x drawn from P_X is encoded as a codeword $K(x)$ which is longer than 25 bits. You may give your answer as a fraction.

4. [1.5 points] Let P be the Markov chain given by the transition diagram



and the initial condition

x	$P_{X_1}(x)$
A	1/2
B	1/2

(a) Compute the probabilities of each of the three initial segments AAA, AAB, and ABA — that is, compute the marginal probabilities (fractions are OK)

$$P(X_1 = \mathtt{A},\ X_2 = \mathtt{A},\ X_3 = \mathtt{A}),$$
$$P(X_1 = \mathtt{A},\ X_2 = \mathtt{A},\ X_3 = \mathtt{B}),$$
$$P(X_1 = \mathtt{A},\ X_2 = \mathtt{B},\ X_3 = \mathtt{A}).$$

(b) Compute the arithmetic codeword for the initial segment ABA.

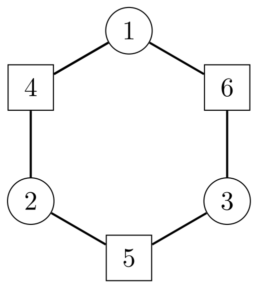
(c) Approximate the probability that $P(X_{100} = \mathtt{A})$.

(d) Compute the entropy rate of the process P.

5. [1 point] A linear code K is given by the generator matrix

$$G^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

This code can also be illustrated by means of the parity check diagram



The content bits are here shown as circles, and the parity bits as squares.

(a) Encode the messages 001 and 111 according to K.

(b) Decode the (noisy) messages 011001, 110011, and 111110.

(c) Find the smallest bit-flip distance (Hamming distance) between two codewords of K.

(d) Find the largest number of bit flips that K is guaranteed to correct.

(e) Find the largest number of bit flips that K is guaranteed to detect.

6. [1 point] Let $\mathcal{X} = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and $\mathcal{Y} = \{2, 3, 5, 7\}$. Let $P_{Y|X}$ be the channel with transition probabilities $P_{Y|X}(y|x) > 0$ if and only if y is a prime factor of x (or, equivalently, $x \equiv 0 \mod y$).

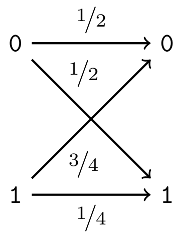
   (a) Give the confusability graph G of the noisy channel $P_{Y|X}$ described above.

   (b) How many messages can be sent perfectly with one use of this channel? Give an explicit description of the encoding and decoding function.

7. [1.5 points] Consider the channel with transition matrix

$$P_{Y|X} = \begin{bmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{bmatrix}$$

and channel diagram



   (a) Write down a formula that expresses, for an arbitrary input distribution over X, the mutual information between X and Y.

   (b) Prove that

$$\frac{\partial}{\partial p} h(a + bp) = b \log_2 \left(\frac{1 - a - bp}{a + bp} \right),$$

   where h is the binary entropy function.

   (c) Use this result to find an approximation to the optimal input distribution for the channel $P_{Y|X}$. Use the cheat sheet from the first page and state explicitly what approximations you use in your derivations.