Examination form

Information Theory

Master of Logic elective course (5314INTH6Y)

Final exam
Date: Friday, March 28, 2014
Time: 9:00-12:00
Form Code: 1481

INSTRUCTIONS
Before starting to work on the examination, please complete the examination form and put it on the corner of your table, along with your **student card** and **photo ID**.

NAME

STUDENT ID NUMBER

SIGNATURE

University of Amsterdam

Faculty of Science

Exam

Information Theory

Master of Logic elective course (5314INTH6Y)

Final exam
Date: Friday, March 28, 2014
Time: 9:00-12:00

Number of pages: 4 (including front page)
Number of questions: 8
Maximum number of points to earn: 9
At each question is indicated how many points it is worth.

BEFORE YOU START

- **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 be put in your coat or bag. Your **coat and bag** should 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 invigilator 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!

1. Let X, Y, Z be *binary* random variables such that $I(X;Y) = 0$ and $I(X;Z) = 0$.
   (a) [$\frac{1}{2}$ points] Does it follow that $I(X;Y,Z) = 0$? If yes, prove it. If no, give a counterexample.
   (b) [$\frac{1}{2}$ points] Does it follow that $I(Y;Z) = 0$? If yes, prove it. If no, give a counterexample.

2. [1 point] Let A, B, C be random variables over alphabet $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ for some integer $n \geq 2$. Let us assume that

$$A = B + C \mod n\,, \tag{1}$$
$$H(B) = \log(n)\,, \tag{2}$$
$$I(A;B) = 0\,. \tag{3}$$

   Show that $I(A;C) = 0$.

3. [1 point] Let A, B, C be random variables such that

$$I(A;B) = 0\,, \tag{4}$$
$$I(A;C|B) = I(A;B|C)\,, \tag{5}$$
$$H(A|BC) = 0\,. \tag{6}$$

   What is the relation between the quantities $H(A)$ and $H(C)$?

4. [1 point] For each value of $m = 0, 1, 2, \ldots$, what is the capacity of the channel consisting of a BSC(ε) together with m symbols which are all transmitted perfectly? The transition matrix of this channel is given by:

$$\left[\begin{array}{cc|c} 1-\varepsilon & \varepsilon & 0\ldots0 \\ \varepsilon & 1-\varepsilon & 0\ldots0 \\ \hline 0 & 0 & \\ \vdots & \vdots & \mathbb{1}_{m\times m} \\ 0 & 0 & \end{array}\right]$$

   where $\mathbb{1}_{m\times m}$ is the m by m identity matrix with 1's on the diagonal and 0's everywhere else.

5. Given a random variable X with the following distribution

x	1	2	3	4	5	6
$P_X(x)$	0.1	0.1	0.3	0.1	0.25	0.15

   (a) [$\frac{1}{2}$ points] Draw a binary Huffman tree which is optimal in average codeword length, and give the corresponding codewords.
   (b) [$\frac{1}{2}$ points] Draw a ternary Huffman tree which is optimal in average codeword length, and give the corresponding codewords.
   (c) [$\frac{1}{2}$ points] Draw a 4-ary Huffman tree which is optimal in average codeword length, and give the corresponding codewords.

6. Both entropy and variance are often used as measures of the "inherent uncertainty" in a distribution, so it is interesting to find out how similar they are. Consider sample space $\mathcal{X} = \{1, 2, \ldots, n\}$ for some $n \geq 2$.

   (a) [$\frac{1}{4}$ points] What distribution $P_X^{\max}$ on $\mathcal{X}$ maximizes the entropy, and what is the entropy $H(P_X^{\max})$?

   (b) [$\frac{1}{4}$ points] What distributions on $\mathcal{X}$ minimize the entropy?

   (c) [$\frac{1}{4}$ points] What distribution $Q_X^{\max}$ on $\mathcal{X}$ maximizes the variance, and what is the variance $\mathrm{Var}[Q_X^{\max}]$?

   (d) [$\frac{1}{4}$ points] What distributions on $\mathcal{X}$ minimize the variance?

   (e) [$\frac{1}{2}$ points] Now let $\mathcal{X}$ be the positive natural numbers. Show that for every $\varepsilon > 0$, no matter how small, and for every finite C, no matter how large, there exists a distribution P_X on $\mathcal{X}$ that has entropy smaller than ε and variance greater than C.

   Hint: For some $\delta > 0$ and $n \geq 2$, consider the distribution $P_X(1) = 1 - \delta$, $P_X(n) = \delta$.

7. In this exercise we consider yet another different entropy notion. Let X and Y be random variables with joint probability distribution P_{XY}. The *collision probability* and the *collision entropy* are respectively defined as

$$\mathrm{Col}(X) := \sum_x P_X(x)^2 \quad \text{and} \quad H_2(X) := -\log \mathrm{Col}(X).$$

   The *conditional collision probability* and the *conditional collision entropy* are respectively defined as

$$\mathrm{Col}(X|Y) := \sum_y P_Y(y)\mathrm{Col}(X|Y = y) \quad \text{and} \quad H_2(X|Y) := -\log \mathrm{Col}(X|Y).$$

   (a) [$\frac{1}{4}$ points] Prove that $H_2(X) \leq H_2(XY)$.

   (b) [$\frac{1}{4}$ points] Prove that $H_2(X|Y) \leq H_2(X)$.

   (c) [$\frac{1}{2}$ points] Prove that

$$0 \leq H_{\min}(X) \leq H_2(X) \leq H(X)$$

   and

$$0 \leq H_{\min}(X|Y) \leq H_2(X|Y) \leq H(X|Y)\,.$$

8. *Zero-error vs non-zero-error Shannon capacity:* Let $P_{Y|X}$ be a discrete memoryless channel with confusability graph G and capacity $C = \max_{P_X} I(X;Y)$.

   (a) [$\frac{1}{3}$ points] Show that $\log(\alpha(G)) \leq C$.

   (b) [$\frac{1}{3}$ points] Show that for any $n \geq 1$, $\log(\alpha(G^{\boxtimes n})) \leq \max_{P_{X^n}} I(X^n;Y^n)$, where the Y^n are obtained by using the channel n times, i.e. $P_{Y^n|X^n}(y^n|x^n) = \Pi_{i=1}^n P_{Y|X}(y_i|x_i)$ for all x^n, y^n.

   (c) [$\frac{1}{3}$ points] Conclude that the zero-error Shannon capacity of G is at most the channel capacity C.