Assignment 3

Q2 2014–2015

Due at 23:59 on 19–12–2014

Requirements: Assignments have to be handed in by each student separately. Please write your name and student number on the top of the first sheet that you hand in. Remember that assignments have to be typeset in English and submitted through the peach system. **Always justify your answers!**

Practice Set 3 contains some information that is relevant to this assignment.

You can score 20 points in total for this assignment; your grade will be the number of points you score divided by two.

Problem 1: [1+1+1+1 points]

In a card trick, a volunteer selects one card among 250 distinct cards. We assume that the volunteer selects the card randomly with a uniform distribution. That is, every card has the same probability of being selected. The magician does not know this choice, but the volunteer will answer a series of questions, where each question has six possible answers. The volunteer always answers truthfully.

1. How much information (measured in bits) does the magician need to obtain from the volunteer to identify the chosen card? Give an exact result (in terms of elementary functions and constants), and a result rounded to three decimal places.
2. What is the maximum amount of information (measured in bits) that the magician can obtain, on average, from the volunteer with each given answer? Give an exact result (in terms of elementary functions and constants), and a result rounded to three decimal places.
3. Give a good lower bound on the minimum number of questions that the magician must ask, on average, to identify the chosen card?
4. What can you say about the minimum number of questions to ask in the worst case?

Problem 2: [1+2+1 points]

According to *Benford's Law*, the first digits of decimal numbers are distributed as

$$P(d) \quad = \quad \log_{10}\left(1+\frac{1}{d}\right) \quad \text{for } 1 \leq d \leq 9$$

You can use the following approximation of this distribution:

```
{
    "1": 0.301,
    "2": 0.176,
    "3": 0.125,
    "4": 0.097,
    "5": 0.079,
```

```
    "6": 0.067,
    "7": 0.058,
    "8": 0.051,
    "9": 0.046
}
```

1. Determine the entropy, measured in bits, of an information source that produces random digits, independently distributed according to Benford's Law. Round your answer to three decimal places.
2. Define a binary encoding of this source that uses, on average, no more that 2.93 bits per symbol.
3. Would it be possible to define an encoding that achieves an average of 2.9 bits per symbol?

Problem 3: [1+1+1+1 points]

The Dutch *Citizen Service Number* (BSN) consists of nine decimal digits. In a BSN, the digits d_i ($i = 1, \ldots, 9$, from left to right), are always such that

$$9d_1 + 8d_2 + 7d_3 + 6d_4 + 5d_5 + 4d_6 + 3d_7 + 2d_8 - d_9$$

is a multiple of 11. Note the minus sign at the end.

1. Which of the following numbers are valid BSNs?

   ```
   222222221
   222222222
   222222322
   222224222
   222252222
   222622222
   227222222
   282222222
   922222222
   ```

2. In a BSN, the fourth digit from the left has become unreadable: `123?56789`. What should that that digit be?
3. Explain why any single-digit error in a BSN will always be *detectable*.
4. Explain why it is not always possible to *correct* a single-digit error in a BSN.

Problem 4: [1+1+2 points]

We use the Hamming $(7, 4)$ error-correcting code, as defined on the slides for Lecture 10.

1. Encode the four source bits `1011`.
2. The word `0011111` was received over a binary symmetric noisy channel. Apply maximum likelihood decoding to determine what code word was most likely sent.

3. We extend every code word with one bit, such that the number of 1-bits is even. The Hamming distance between two code words is then always even and, therefore, at least 4. How many errors can be *detected* by this new code?

Problem 5: [1+1+2 points]

This problem partly involves the use of GPG (*GNU Privacy Guard*). You should first download the *public key for this course*, and import it into your local keyring:

```
gpg --import pubkey-2IS80.txt
```

1. Consider these four almost identical files:

   - `LBUI-1.txt`
   - `LBUI-2.txt`
   - `LBUI-3.txt`
   - `LBUI-4.txt`

   and these two detached signatures:

   - `signature-A-1415.asc`
   - `signature-B-1415.asc`

   Determine which file was signed with which signature. You verify that detached signature `signature.asc` is valid for file `file.txt` by

   ```
   gpg --verify signature.asc file.txt
   ```

2. When you have completed the PDF with your answers for submission to peach, encrypt that PDF using the *public key for this course*, and submit the result to peach, *together with your unencrypted PDF*. Suppose that your answers are in the file `answers.pdf`, then you encrypt it with

   ```
   gpg -r 2is80 -a -o answers.txt --encrypt answers.pdf
   ```

   The option `-r 2is80` indicates the recipient (assuming you do not have other public keys on your keyring that contain `2is80`); `-a` (for ASCII armor) indicates plain text output. Submit both `answers.pdf` and `answers.txt`.

3. Explain in your own words what the danger is when you encrypt with a block cipher *without using cipher block chaining*.