Problem sheet 1 2020

Ex. 1

Suppose a binary repetition code of length 5 is used for a binary symmetric channel with (symbol error) crossover probability p. Calculate the word error probability, i.e. the probability that a word of length 5 is incorrectly received, even after error correction. Evaluate this probability if p = 0.1.

Ex. 2

We want to find the best possible 3-ary (n, M, d) code, where q = 3, n = 3 is the word length, M is the number of codewords, and d = 2 is the minimum distance of the code. What is the largest M one can use?

- a) Show that a 3-ary (3, M, 2)-code must have $M \leq 9$.
- b) Show that a 3-ary (3,9,2)-code exists. (Hint: find three codewords starting with 0, and three codewords starting with 1, and three codewords starting with 2).

Ex. 3

Let $E_n \subset F_2^n$ denote the set of all vectors with even weights. Deduce that E_n is the code that is obtained by adding a parity check to the code $C = F_2^{n-1}$. Deduce that E_n is an $(n, 2^{n-1}, 2)$ -code.

Ex. 4

Prove that $A_q(3,2) = q^2$.

Ex. 5

Show: If a binary (n, M, d)-code exists, with d even, then there also exists a binary (n, M, d)-code in which all codewords haven even weight.

Ex. 6

Each (properly published) book gets a unique ISBN number (international standard book number). This is a 10-digit codeword. The first digit stands for the country/language, the next few digits for the publisher. Then some digits for a number assigned by the publisher, the very last digit is a checksum. (A large publisher gets a short publisher identification and can thus use more digits for its own books, a small publisher gets a longer publisher identification. This alone leads to interesting questions but we leave these aside.)

For example, the recommended text book by Ray Hill has the number ISBN 0-19-853804-9

ISBN 0-19-853803-0 (for the paperback edition).

Here the first 0 stands for english, the 19 for Oxford University Press.

Let $x_1 x_2 \cdots x_{10}$ be the ISBN number (codeword). The check bit x_{10} is chosen such that the whole codeword satisfies $\sum_{i=1}^{10} ix_i \equiv 0 \mod 11$.

a) Show that $x_{10} = \sum_{i=1}^{9} ix_i \mod 11$.

Note that the last symbol can be any of 11 eleven values. So, one uses in addition to $0, 1, \ldots, 9$ the symbol X = 10.

- b) Show that this code can be used in the following way: To detect any single error and to detect a double error created by the transposition of two digits (example 152784 ↔ 158724).
 Would this also work, if you use a similar code mod 15 instead of mod 11?
- c) Can this method be used to correct one single error?
- d) Discuss the advantages of this method for the practical use (to order books in a bookshop etc.).
- e) What is the minimum distance of any two ISBN numbers?
- f) Consider a different code C_2 , where one uses as before 10 digits but does not use a weighted sum, but $\sum_{i=1}^{10} x_i \equiv 0 \mod 11$. What would be the disadvantage, compared with the ISBN code?
- g) The 10-digit ISBN code has meanwhile been incorporated into the 13-digit EAN-13 code. For details see this webpage https://www.arndt-bruenner. de/mathe/scripts/pruefziffern.htm

Ex. 7

Design a guessing game as follows:

Tom thinks of an integer $w \in \{0, ..., 15\}$. Scarlett aksks four questions, which Tom correctly answers with yes or no, and Scarlett then tells Tom the number he had chosen.

In a second round Tom is allowed to lie once. Again he thinks of a number $w \in \{0, ..., 15\}$. Which questions (strategy) should Scarlett use, and how many questions does she need to find Tom's number? (In other words, give a best possible strategy for Scarlett. Describe it, if possible, in a way that involves coding (which code?))

The strategy can (possibly?) be implemented as follows:

Scarlett's questions are equivalent to "Is $w \in A_i$, where $A_i \subset \{0, \dots, 15\}$?" In this way, Scarlett can ask all questions simultaneously, and in particular the answers to the first questions do not influence the later *questions*. (Has this last property any potential application?)