MA210 - Class 9

CODES
- Practical reasons
- How to actually apply it

FORMAL DEFINITION ① $C \subseteq \{0,1\}^n$

② d-error-detecting. $\forall c \in C\ \forall x \in \{0,1\}^n$, $d_H(x,c) \leq d \wedge x \neq c \Rightarrow x \notin C$. $\Leftrightarrow \delta(C) \geq d+1$, $|C| \geq 2$

③ d-error-correcting. $\forall c \in C\ \forall x \in \{0,1\}^n$, $d_H(x,c) \leq d$, then the nearest C-neigh. of x exists and is c.
$\Leftrightarrow \delta(C) \geq 2d+1$

STANDARD-PARITY CHECK IMPLEMENTATION.

Follow the referenced link.

3. You are asked to design a binary code for a certain application. For technical reasons, the code must satisfy the following conditions:

(i) at most three consecutive 1s are allowed in every codeword;

(ii) every 0 must be preceded or followed by another 0. (In other words, no codeword should start like 01 . . . , end like . . . 10, or look like . . . 101 . . . .)

(a) Give all possible codewords of length 5.

You must design a code of length 125. So let C be the code formed by taking all words of length 125 that satisfy the two conditions above.

(b) Show that for every codeword $\mathbf{x} \in C$, the weight $w(\mathbf{x})$ of $\mathbf{x}$ satisfies $w(\mathbf{x}) \leq 75$.

⇝ MOST COMMON MISTAKE: "the way of reaching maximal weight is 11100 11100 . . . ". This is just false.
"To maximize the weight we want to maximise the # of 111". Also FALSE

⇝ RIGHT IDEA: Pigeonhole principle

(b) Suppose that C_1 and C_2 are two linear codes of length n.

(i) Show that $C_1 \cap C_2$ is a linear code.

(ii) Is $C_1 \cup C_2$ a linear code in general? Justify your answer by a proof or a counterexample.

Def (LINEAR CODE) A code $C \subseteq \{0,1\}^n$ is linear if $\forall x,y \in C$, $x+y \in C$.

IDEA a) Take $x,y \in C_1 \cap C_2$. Then $x+y \in C_1$ bc C_1 is linear and $x+y \in C_2$ bc C_2 is linear.

b) Go for EASY examples. $C_1 = \{00, 01\}$, $C_2 = \{00, 10\}$.
$C_1 \cup C_2 = \{00, 01, 10\}$ which is NOT linear since $10 + 01 \notin C_1 \cup C_2$.

5. (a) Define the Hamming distance $d_H(\boldsymbol{x}, \boldsymbol{y})$ of two words $\boldsymbol{x}, \boldsymbol{y}$ in $\{0,1\}^n$.

(b) Let $\boldsymbol{x}$ be a word in $\{0,1\}^n$ and let k be an integer such that $0 \leq k \leq n$.

(i) How many words $\boldsymbol{y}$ in $\{0,1\}^n$ satisfy $d_H(\boldsymbol{x}, \boldsymbol{y}) = k$?

(ii) How many words $\boldsymbol{y}$ in $\{0,1\}^n$ satisfy $d_H(\boldsymbol{x}, \boldsymbol{y}) \leq k$?

Justify your answers.

(c) Suppose that C is a 2-error-correcting code. Explain why

$$|C| \leq \frac{2^n}{1 + n + \binom{n}{2}}.$$

(d) Prove that there is no code C of length 8 such that $|C| \geq 7$ and $\delta(C) = 5$.

$\rightsquigarrow$ IDEAS $\binom{n}{k}$ b.i) Fix x. Now to get y s.t. $d_H(x,y) = k$, they must differ in EXACTLY k places. Also, these k positions uniquely identify y (since x is fixed).
There are $\binom{n}{k}$ ways of choosing k positions among n.

c) 2 error correcting $\Rightarrow \forall x, y \in C$, $N^2(x) \cap N^2(y) = \emptyset$.
Therefore we have

$$|\{0,1\}^n| \geq \sum_{x \in C} |N^2(c)| = |C| \cdot \left[\binom{n}{2} + n + 1\right]$$

d) from C.

(a) Let C be the following code of length 6:

$$C = \{\underbrace{000000}_{a}, \underbrace{010101}_{b}, \underbrace{101010}_{c}, \underbrace{111000}_{d}, \underbrace{000111}_{e}, \underbrace{111111}_{f}\}.$$

(i) How many errors can C detect? And how many errors can C correct? Justify your answer.

(ii) In order to improve the error-correcting properties of the code C, you add a 7-th bit at the end of each word, which must of course be 0 or 1.

Describe all ways that a 7-th bit can be added to each word so that the resulting code can correct more errors than C can. Justify your answer.

i) $d(c,d) = d(e,b) = 2$

(b) All codes in this question are binary codes.

For a word $\bar{\boldsymbol{x}} \in \{0,1\}^m$, $\bar{\boldsymbol{x}} = x_1 x_2 \ldots x_m$, and a word $\bar{\boldsymbol{y}} \in \{0,1\}^n$, $\bar{\boldsymbol{y}} = y_1 y_2 \ldots y_n$, define the *concatenation* $\bar{\boldsymbol{x}} \bullet \bar{\boldsymbol{y}}$ as the word of length $m+n$ by putting the words behind each other: $\bar{\boldsymbol{x}} \bullet \bar{\boldsymbol{y}} = x_1 x_2 \ldots x_m y_1 y_2 \ldots y_n$.

Similarly, for a code C_1 of length m and a code C_2 of length n, the *concatenation* $C_1 \bullet C_2$ is defined as $C_1 \bullet C_2 = \{\, \bar{\boldsymbol{x}} \bullet \bar{\boldsymbol{y}} \mid \bar{\boldsymbol{x}} \in C_1,\ \bar{\boldsymbol{y}} \in C_2 \,\}$.

(i) For $\bar{\boldsymbol{x}}_1, \bar{\boldsymbol{x}}_2 \in \{0,1\}^m$ and $\bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2 \in \{0,1\}^n$, find an expression of $d_H(\bar{\boldsymbol{x}}_1 \bullet \bar{\boldsymbol{y}}_1, \bar{\boldsymbol{x}}_2 \bullet \bar{\boldsymbol{y}}_2)$ in terms of $d_H(\bar{\boldsymbol{x}}_1, \bar{\boldsymbol{x}}_2)$ and $d_H(\bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2)$.

(ii) Is it the case that for two binary codes C_1, C_2 we have $\delta(C_1 \bullet C_2) = \delta(C_1) + \delta(C_2)$? Justify your answer.

Suppose C_1 and C_2 are both linear codes.

(iii) Show that $C_1 \bullet C_2$ is also a linear code.