ERROR-CORRECTING CODES

MARCH 3, 2024

Given two sequences a and b of 0 and 1, both of length n, define the distance between them as the number of positions in which they differ. E.g., distance between 10001 and 10010 is 2.

For brevity, we will call a sequence of 0 and 1 a *word*.

1. Given a word a of length 7, how many words (of same length) are there which are at distance at most 1 from a (including a itself)?

2. Is it possible to choose 20 different words of length 7 such that any two of them differ in at least 3 places? What about 10 words?[Hint: this is equivalent to asking that "balls of radius 1" around these words do not intersect.]

3. Show that one can choose 16 different words $C_1, \ldots, C_{16}$ of length 7 such that any two of them differ at at least 3 positions.

4. An error correcting code is a way of assigning to every sequence a of k bits a longer sequence C_a of n bits ($n > k$) in such a way that if you transmit C_a over some noisy communication line, and one bit is corruputed, the recipient can still uniquely restore C_a. The simplest such way is just repeating every bit of a three times, but it is wasteful. [To be precise, the above definition is of the code that can correct a single bit error.]

Use the problems above to construct an error-correcting code which allows you to encode a 4-bit sequence by a 7-bit sequence, and yet correct a single bit error.