Chapter 5: Error correcting codes

5.1 Definitions and some bounds on codebook size

We have alphabet Σ with $0 \in \Sigma$.

Def: The Hamming weight of a string $x^n \in \Sigma^n$ is as $|\{i : x_i \neq 0\}|$. The Hamming distance between $x^n, y^n \in \Sigma^n$ is

$$\delta(x^n, y^n) = |\{i : x_i \neq y_i\}|$$

Def. An error correction code C of length n over Σ is a subset of Σ^n. C is called a codebook.

- C is a binary code if $\Sigma = \{0, 1\}$
- A binary code C is a linear code if it is a subspace of $\{0,1\}^n$, i.e. for any $c, c' \in C$, $c \oplus c'$ is also in C. It always contains 0^n.
- The size of the codebook is denoted $|C|$.
- The rate of the code is

$$R(C) = \frac{\log |C|}{n \log |\Sigma|}$$

- The minimal distance of C is

$$d(C) = \min \delta(c, c')$$

$$d(C) = \min_{\substack{c,c' \in C \\ c \neq c'}} d(c,c')$$

Example

Take strings $\{0,1\}^n$ and add a parity bit

$$x_{n+1} = \bigoplus_{i=1}^{n} x_i .$$

$\Rightarrow$ binary code, $|C| = 2^n$, $R(C) = \frac{n}{n+1}$

$$d(C) = 2$$

Consider a binary code with minimum distance d

- Detect up to $d-1$ bit flip errors
- correct up to $\lfloor \frac{d-1}{2} \rfloor$ bit flip errors
- correct up to $d-1$ erasures

Lemma: Hamming bound:

$$|C| \leq \frac{2^n}{\sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i}}$$

(e.g. for $d=3$ we have $|C| \leq \frac{2^n}{n+1}$)

Proof: For every $c \in C$, define its neighborhood $N(c,r)$ as all the strings with distance at most r from c.

Set $r = \lfloor \frac{d-1}{2} \rfloor$, then $N(c,r) \cap N(c',r) = \emptyset$ for all $c, c' \in C$, $c \neq c'$ $\quad (*)$

Then $2^n \geq \left| \bigcup_{c \in C} N(c,r) \right| = \sum_{c \in C} |N(c,r)|$

uses (*)

$= |C| \sum_{i=0}^{r} \binom{n}{i}$ $\square$

A perfect code satisfies $\bigcup_{c \in C} N(c,r) = \{0,1\}^n$.

Lemma: Singleton bound. Assume $|\Sigma| = q$, then $|C| \leq q^{n-d+1}$.