CS588 Notes on Entropy and Perfect Ciphers

Entropy: Amount of information in a message

$H(M) = -\Sigma P(M_i) \log P(M_i)$ over all possible messages M_i

If there are n equally probable messages with a binary alphabet,

$H(M) = \log_2 n$

Absolute Rate (R): how much information can be encoded

$R = \log_2 Z$ (Z=size of alphabet)

Actual Rate (r): how much information can be encoded

$r = H(M) / N$

number of possible N-letter messages

Redundancy (D):

$D = R - r$

In English, $D \approx 1 - .28 = .72$ letters/letter

Entropy of cryptosystem: (K = number of possible keys)

$H(K) = \log_{\text{Alphabet Size}} K$ if all keys equally likely

Unicity distance:

$U = H(K)/D$

Perfect Cipher:

$\forall$ i, j: $P(M_i|C_j) = P(M_i)$

A cipher is perfect iff:

$\forall M, C \qquad P(C \mid M) = P(C)$

Or, equivalently:

$\forall M, C \qquad P(M \mid C) = P(M)$

Perfect Cipher Keyspace Theorem: If a cipher is perfect, there must be at least as many keys (l) are there are possible messages (n).

Proof:

Suppose there is a perfect cipher with $l < n$. (More messages than keys.) Let C_0 be some ciphertext with $p(C_0) > 0$. There exist

m messages M such that $M = D_K(C_0)$

$n - m$ messages M_0 such that $M_0 \neq D_K(C_0)$

We know $1 \leq m \leq l < n$ so $n - m > 0$ and there is at least one message M_0.

Consider the message M_0 where $M_0 \neq D_K(C_0)$ for any K.

So,

$p(C_0 \mid M_0) = 0.$

In a perfect cipher,

$p(C_0 \mid M_0) = p(C_0) > 0.$

Hence, by contradiction all perfect ciphers must have $l \geq n$.