Entropy

Entropy

- Entropy is a measure of uncertainty and of how much information can be stored in a unit, so that we can accurately represent all outcomes of an event.

- Definition: Suppose X is a discrete random variable which takes on values from a finite set X. Then the entropy of the random variable X is defined to be the quantity

$$H(X) = -\sum_{x \in X} \Pr[x] \log_2 \Pr[x]$$

Properties of Entropy

- Theorem 1:

Suppose X is a random variable having a probability distribution which takes as the values p_1, p_2, ….., p_n, where $p_i > 0$, $1 \le i \le n$. Then

$$H(X) \le \log_2 n, \text{ with equality if and only if}$$
$$p_i = 1/n\ ,\ 1 \le i \le n.$$

- Theorem 2:

$H(X,Y) \le H(X) + H(Y)$, with equality if and only if X and Y are independent random variables.

Conditional Entropy

- Definition: Suppose X and Y are two random variables. Then for any fixed value y of Y, we get a (conditional) probability distribution on X; we denote the associated random variable by X|y.

$$H(X|y) = -\sum_{x \in X} \Pr[x|y] \log_2 \Pr[x|y]$$

we define the conditional entropy, denoted H(X|Y), to be the weighted average (with respect to the the probabilities Pr[y]) of the entropies H(X|y) over all possible values of y. It is computes as

$$H(X|Y) = -\sum_{y} \sum_{x} \Pr[y] \Pr[x|y] \log_2 \Pr[x|y]$$

- The conditional entropy measures the average amount of information about X that is revealed by Y

Conditional Entropy contd.

Theorem 3: $H(X,Y) = H(Y) + H(X|Y)$

Corollary 1: $H(X,Y) \leq H(X)$, with equality if and only if X and Y are independent.

Spurious Keys and Unicity Distance

Relationship among the entropies of th components of a cryptosystem

Theorem: Let (P, C, K, E, D) be a cryptosystem. Then

$$H(K|C) = H(K) + H(P) - H(C)$$

Hint:

$H(K,P,C) = H(K,P) = H(K) + H(P)$

$H(K,P,C) = H(K,C)$, since $H(P|K,C) = 0$, $x = d_k(y)$

Compute $H(K|C) = H(K|C) - H(C)$

$= H(K,P,C) - H(C)$

$= H(K) + H(P) - H(C)$

Spurious Keys

- Suppose we have a cryptosystem and a plaintext x encrypted with a key k resulting in ciphertext y. Knowing only the ciphertext y, how can we determine the key?
- Many keys may remain, only one of which is the correct. Keys which are possible but incorrect are called spurious keys.
- Example:

  Ciphertext (shift cipher) – WNAJW

  with k = 5, meaningful plaintext – river

  with k = 22, meaningful plaintext – arena

Spurious Keys contd.

- How much information can a language store?

Answer: We measure this by HL, the entropy per letter of a natural language. This is the average information per letter in a meaningful string of text.

H(P) - the entropy of the random variable associated with the plaintexts. $H(P^n)$ - the entropy of the random variable representing plaintexts of length n.

Definition: Suppose L is a natural language. The entropy of L is defined to be the quantity $HL = \lim_{n\to\infty} H(P^n)/n$, where P^n is the random variable that has its probability distribution that of all plaintexts of length n. We also define the redundancy of L to be given by $RL = 1 - (HL / \log_2 |P|)$.

RL = 0.75 for the English language, so the English language is 75% redundant!

Spurious Keys

- Theorem:

Suppose (P, C, K, E, D) is a cryptosystem where $|C| = |P|$ and keys are chosen with the same probability. Then given a ciphertext of length n, the expected number of spurious keys satisfies $s_n \geq (|K| / (|P|^{n R_L}))- 1$

Here, R_L denotes the redundancy of the language.

Unicity Distance

- The unicity distance of a cryptosystem is the the average size of ciphertext (value of n) at which the expected number of spurious keys becomes zero.
- Using our previous theorem, we get an estimate for the unicity distance as

$$n_o = \log_2 |K| / (R_L \log_2 |P|)$$

The average amount of ciphertext required for an third party to be able to uniquely compute the key, given enough computing time.