# Information Theory and Coding - Prof. Emere Telatar

Jean-Baptiste Cordonnier, Sebastien Speierer, Thomas Batschelet

October 4, 2017

## 1 Data compression

Given an alphabet $\mathcal{U}$ (e.g. $\mathcal{U} = \{a, ..., z, A, ..., Z, ...\}$), we want to assign binary sequences to elements of $\mathcal{U}$, i.e.

$$e : \mathcal{U} \to 0, 1^* = \{\emptyset, 0, 1, 00, 01, ...\}$$

For $\mathcal{X}$ a set

$$\mathcal{X}^n \equiv \{(x_0...x_n), x_i \in \mathcal{X}\}$$
$$\mathcal{X}^* \equiv \bigcup_{n \geq 0} \mathcal{X}^n$$

**Definition 1.**
A code $\mathcal{C}$ is called *singular* if

$$\exists (u, v) \in \mathcal{U}^2, u \neq v \quad s.t. \quad C(u) = C(v)$$

Non singular code is defined as opposite

**Definition 2.**
A code $\mathcal{C}$ is called *uniquily decodable* if

$$\forall u_1, ..., u_n, v_1, ..., v_n \in \mathcal{U}^* \quad s.t. \quad u_1, ..., u_n \neq v_1, ..., v_n$$

we have

$$\mathcal{C}(u_1)\mathcal{C}(u_n) \neq \mathcal{C}(v_1)\mathcal{C}(v_n)$$

i.e, $\mathcal{C}^*$ is non-singular

**Definition 3.**
Suppose $\mathcal{C} : \mathcal{U} \to \{0, 1\}^*$ and $\mathcal{D} : \mathcal{V} \to \{0, 1\}^*$ we can define

$$\mathcal{C} \times \mathcal{D} : \mathcal{U} \times \mathcal{V} \to \{0, 1\}^*$$

as

$$(\mathcal{C} \times \mathcal{D})(u, v) \to \mathcal{C}(u)\mathcal{D}(v)$$

**Definition 4.**
Given $\mathcal{C} : \mathcal{U} \to \{0,1\}^*$, define

$$\mathcal{C}^* : \mathcal{U}^* \to \{0,1\}^*$$

as

$$\mathcal{C}^*(u_1, u_n) = \mathcal{C}(u_1)...\mathcal{C}(u_n)$$

#### 1.0.1 Markov chains

For a Markov Chaine $A \to B \to C \to D$, the joint probability distribution of the RVs should be $p(a)p(b|a)p(c|b)p(d|c)$

- The reverse of a MC is a MC

**Kraft-sum** **Definition:** The Kraftsum of a code $C$ is $KS(C) = \sum_u 2^{-|C(u)|}$

- if $C$ is prefix free then $KS(C) \leq 1$
- if $C$ is non singular, then $KS(C) \leq 1 + \min_u |C(u)|$
- $KS(C^n) = KS(C)^n$

**Theorem:** for any $U$ and associated $p(u)$ there exists a prefix free code $C$ s.t.

$$E[|C(U)|] < 1 + \sum_{u \in U} p(u) \log \frac{1}{p(u)}$$

**Theorem:** if $KS(C) \leq 1$ then there exists a prefix free code $C'$ such that $|C(u)| = |C'(u)|$ for all $u$
**Corollar:** if $C$ is uniquely decodable, then there exists $C'$ that is prefix free with the same word lengths

**Entropy** **Definition:** the entropy of a random variable $U$ is

$$H(U) = \sum_{u \in U} p(u) \log \frac{1}{p(u)} = E_U \left[ \log \frac{1}{p(u)} \right]$$

**Theorem:** if $C$ is uniquely decodable then $E[|C(U)|] \geq H(U)$

**Properties of optimal prefix free codes**

1. $p(u) < p(v) \to |u| \geq |v|$
2. The two longest codewords have the same length
3. The 2 least probable letters are assigned codewords that differ in the last bit

#### 1.0.2 Hoffman algorithm

- Combine the 2 least likely symbols
- Sum their probability and assign it a new fictive symbol
- Repeat