Information Theory and Coding - Prof. Emere Telatar

Jean-Baptiste Cordonnier, Sebastien Speierer

October 4, 2017

0.0.1 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

0.0.2 Hoffman algorithm

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