Wolfgang K. Seiler
Tel. 2515

B6, B4.18

October 5, 2018

Mathematics and Information, Exercise sheet 5

Problem 1: (*8 points*)
Let $\mathcal{X} = (X_t)_{t\in\mathbb{N}}$ be a sequence of random variables with values in an alphabet A with m elements, and for $a \in A^n$ let $p(a) = p(X_1 = a_1, \dots, X_n = a_n)$. The *uncertainty* of $\mathcal{X}$ is

$$\mathcal{U} = \lim_{n\to\infty} -\frac{\sum\limits_{a\in A^n} p(a)\log_2 p(a)}{n\log_2 m},$$

and the *redundancy* is $\mathcal{R} = 1 - \mathcal{U}$.

a) Show that $0 \le \mathcal{R} \le 1$.

b) If we model the english language by a sequence of independent random variables, its entropy per letter is about 4.22295. Compute the redundancy of this process!

c) If we model it by a Markov chain, the entropy rate is about 3.62773. What's the redundancy now?

d) It is commonly believed that the redundancy of the english language is about 70%. What does this mean for the entropy per letter?

Problem 2: (*5 points*)
For each of the three processes in problem 1, compute the probability that a given sequence of hundred letters lies in the set O of frequent blocks!

Problem 3: (*5 points*)
Let $\mathcal{X} = (X_t)_{t\in\mathbb{N}}$ be e sequence of independent random variables with values in $A = \{a, b\}$, where each X_i takes a with a probability of 2/3. Determine for $\varepsilon = \frac{1}{2}$ and $\varepsilon = \frac{1}{4}$ which percentage of the elements of A^{10} lies in A_ε^{10}!

Problem 4: (*2 points*)
Show that for every positive real number a, $\lim\limits_{k\to\infty} \sqrt[k]{ak} = 1$!

Can be submittel till Thursday, October 11, 2018 at 11.55h