Mathematical methods in communication

Proof of Kraft's Inequality

Author: Ziv Goldfeld *Date: 21.04.2011*

I. Alternative Proof of Kraft's Inequality

Theorem 1 (Kraft's Inequality)

(i) For any prefix code $\{c_i\}_{i\geq 1}$, with lengths $\{l_i\}_{i\geq 1}$ we have:

$$\sum_i 2^{-l_i} \leq 1 \tag{1}$$

(ii) Conversly if $\{l_i\}$ satisfy (1), then there exists a prefix code with these lengths.

Remark 1 The following proof of Kraft's inequality is preferable compared to the previous proof that was presented because it doesn't demand a finite set of codewords or lengths.

Proof: of (i):

Let $\{c_i\}$ be a prefix code, where c_i is a codeword of length $l_i = |c_i|$. We define a function $f : c_i \longrightarrow [0,1]$ that calculates the decimal value of c_i, by:

$$f(c_i) = \sum_{j=1}^{l_i} c_{i,j} \cdot 2^{-j}$$

For future reference, we inspect the interval $\left[f(c_i), f(c_i) + 2^{-l_i}\right)$. Note that:

1. $0 \leq f(c_i) \leq 1$.
2. $f(c_i 000\ldots 0) = f(c_i)$. i.e. adding zeroes at the end of the codeword does not change the value of $f(c_i)$.
3. $f(c_i 111\ldots) = f(c_i) + \sum_{j=1}^{\infty} 2^{(-l_i + j)} = f(c_i) + 2^{-l_i}$. This follows from:

$$q^n - 1^n = \left(1 + q + q^2 + \ldots + q^{n-1}\right)(q-1),$$

which gives

$$1 + q + q^2 + q^3 + \ldots + q^n = \frac{q^{n+1} - 1}{q - 1}.$$

Without loss of generality, we may assume that the $\{c_i\}$ are arranged in an increasing lexicographic order, which means that $f(c_i) \leq f(c_k)$ for all $i \leq k$.

Since c_i is a prefix code we have:

$$f(c_{i+1}) \geq f(c_i) + 2^{-l_i} \tag{2}$$

Thus, we get that the intervals $\left[f(c_i), f(c_i) + 2^{-l_i}\right)$ are pairwise disjoint.

By recurrent use of Inequality (2) we obtain:

$$f(c_m) \geq \sum_{i=1}^{m} 2^{-l_i}$$

Since, by definition $f(c_m) \leq 1$, this proves the first part of the theorem, i.e.

$$\sum_{i=1}^{m} 2^{-l_i} \leq 1$$

■

We have seen that a necessary condition for a code $\{c_i\}$ to be prefix is that the intervals$\left[f(c_i), f(c_i) + 2^{-l_i}\right)$ are pairwaise disjoint. The proof of the second part of the theorem is based upon the claim that this condition is also sufficient:

Lemma 1 Given a code $\{c_i\}$ such that the intervals $\left[f(c_i), f(c_i) + 2^{-l_i}\right)$ are disjoint, the code is prefix.

Remark 2 In the following proof we use the fact that in order to prove $A \Rightarrow B$ one can show that $B^c \Rightarrow A^c$ (i.e. not B $\Rightarrow$not A).

Proof: We conversly assume that the code $\{c_i\}$ is not prefix. If it is so, we can find two codewords c_m and c_n (without loss of generality we assume $m > n$ thus $l_m > l_n$), for which the first $|c_n|$ bits of c_m are identical to the bits of c_n. In this case:

$$f(c_m) = \sum_{j=1}^{l_m} c_{m,j} \cdot 2^{-j} = \sum_{j=1}^{l_n} c_{n,j} \cdot 2^{-j} + \sum_{j=l_n+1}^{l_m} c_{m,j} \cdot 2^{-j} < f(c_n) + 2^{-l_n}$$

So we get that $f(c_m) < f(c_n) + 2^{-l_n}$, contradictly to the fact that the intervals $\left[f(c_n), f(c_n) + 2^{-l_n}\right)$ and $\left[f(c_m), f(c_m) + 2^{-l_m}\right)$ are pairwise disjoint. Thus the code is prefix. $\blacksquare$

Proof: of (ii):

Assume that the lengths $\{l_i\}$ are given and satisfy Kraft's inequality (1). We prove that we can find a prefix code with the given lengths. Without loss of generality, assume that $l_1 \leq l_2 \leq ....$ We define the word c_i to be the inverse image under the mapping f of the number $\sum_{j=1}^{i-1} 2^{-l_j}$, i.e. c_i is the only word (up to addition of zeroes from the right) such that the equality

$$f(c_i) = \sum_{j=1}^{i-1} 2^{-l_j}$$

holds.

To calculate c_i we use the function $f^{-1} : [0, 1] \longrightarrow c_i$. In order to justify that use we first show that $0 < f(c_i) \leq 1$.

From the structure of $f(c_i)$ it is easy to see that $f(c_i) > 0$ for every i. Moreover, using the assumption of the theorem (i.e. inequality (1)) we get that

$$f(c_i) = \sum_{j=1}^{i-1} 2^{-l_j} \leq 1$$

for every i. Thus we get that $0 < f(c_i) \leq 1$.

Next show that the length of every codeword c_i that is built this way is indeed no longer than l_i.

Again from the structure of $f(c_i)$, it is simple to see that the maximal number of bits needed for the codeword c_i is l_{i-1} bits. Because we assume that the lengths are arranged by rising order (i.e. $l_{i-1} \leq l_i$ for every i), the length of each codeword c_i cannot be longer than $|c_i| = l_i$. If it is shorter, we add zeroes from the right up to the wanted length.

To complete the proof, it is enough to show that the intervals

$$I_i = \left[f(c_i), f(c_i) + 2^{-l_i}\right) = \left[\sum_{j=1}^{i-1} 2^{-l_j}, \sum_{j=1}^{i} 2^{-l_j}\right)$$

are pairwise disjoint and finally use Lemma 1.

Since by definition, $f\left(c_i\right)$ increases as i increases and the right border of the interval I_i is the left border of the interval I_{i+1} , the intervals $\{I_i\}$ are pairwise disjoint, which concludes the proof. $\blacksquare$