Theorem 4.4 (Kraft Inequality) Let $\mathcal{C}$ be a D-ary source code, and let $l_1, l_2, \cdots, l_m$ be the lengths of the codewords. If $\mathcal{C}$ is uniquely decodable, then

$$\sum_{k=1}^{m} D^{-l_k} \leq 1. \tag{1}$$

Proof

1. Without loss of generality, assume

$$l_1 \leq l_2 \leq \cdots \leq l_m.$$

2. Let N be an arbitrary positive integer, and consider

$$\begin{aligned} &\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \\ &= \sum_{k_1=1}^{m} \sum_{k_2=1}^{m} \cdots \sum_{k_N=1}^{m} D^{-(l_{k_1}+l_{k_2}+\cdots+l_{k_N})}. \end{aligned}$$

3. By collecting terms of the same degree on the RHS, we write

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N = \sum_{i=1}^{Nl_m} A_i D^{-i}. \tag{2}$$

where A_i is the coefficient of D^{-i} on the LHS.

3. Now observe that A_i gives the total number of sequences of N codewords with a total length of i code symbols (see Example 4.5).

4. Since the code is uniquely decodable, these code sequences must be distinct, and therefore

$$A_i \leq D^i \tag{3}$$

because there are D^i distinct sequences of i code symbols.

5. Substituting (3) into (2), we have

$$\left(\sum_{k=1}^{m} D^{-l_k}\right)^N \leq \sum_{i=1}^{Nl_m} D^i D^{-i} = \sum_{i=1}^{Nl_m} 1 = Nl_m,$$

or

$$\sum_{k=1}^{m} D^{-l_k} \leq (Nl_m)^{1/N}.$$

Since this inequality holds for any N, upon letting $N \to \infty$, we obtain (1), completing the proof.