Capacity of weakly \((d, k)\)-constrained sequences

Kees A. Schouhamer Immink
Institute for Experimental Mathematics, Ellernstrasse 29,
45326 Essen, Germany.
immink@exp-math.uni-essen.de

Augustus J.E.M. Janssen
Philips Research Laboratories,
WY 81, Prof. Holstlaan 4, 5655 AA Eindhoven, The Netherlands.
a.j.e.m.janssen@philips.com

Abstract — In the presentation we find an analytic expression for the maximum of the normalized entropy
\[ -\sum_{i\in T} p_i \ln p_i / \sum_{i\in T} p_i, \]
where the set \( T \) is the disjoint union of sets \( S_n \) of positive integers that are assigned probabilities \( P_n, \sum_n P_n = 1 \). This result is applied to the computation of the capacity of weakly \((d, k)\)-constrained sequences that are allowed to violate the \((d, k)\)-constraint with small probability.

I. PROBLEM DESCRIPTION AND RESULTS

Let \( T \) be a set of positive integers, and assume that \( T \) is the disjoint union of a (finite or infinite) number of non-empty sets \( S_n, n \in M \). Also assume that there are given numbers \( P_n \geq 0, n \in M, \) with \( \sum_n P_n = 1 \). We show the following result.

Theorem: The maximum of
\[ H := -\sum_{i\in T} p_i \ln p_i \]
(\( \ln \) : natural logarithm) under the constraints that \( p_i \geq 0, \sum_{i\in S_n} p_i = P_n, n \in M \), equals \( z_0 \), where \( z_0 > 0 \) is the unique solution of the equation
\[ -\sum_{n\in M} P_n \ln Q_n(z) = -\sum_{n\in M} P_n \ln P_n \]
with \( z > 0 \)
\[ Q_n(z) := \sum_{i\in S_n} e^{-iz}, \quad n \in M. \]

Moreover, the optimal \( p_i \) are given by
\[ p_i = \frac{P_n}{Q_n(z_0)} e^{-iz_0}, \quad i \in S_n, n \in M, \]
and for these \( p_i \) we have that
\[ \sum_{i\in T} ip_i = \frac{d}{dz} \left[ -\sum_{n\in M} P_n \ln Q_n(z) \right] (z_0). \]

As an application of this result we consider weakly constrained \((d, k)\) sequences [1]. A binary \((d, k)\)-constrained sequence has by definition at least \( d \) and at most \( k \) 'zeros' between consecutive 'ones'. Weakly constrained codes produce sequences that violate the specified constraints with a small probability. It is argued that if the channel is not free of errors, it is pointless to feed the channel with perfectly constrained sequences. A \((d, k)\)-constrained sequence can be thought to be composed of 'phrases' \( 1^i \), \( d \leq i \leq k \), where \( 0^i \) means a series of \( i \) 'zeros'. In order to compute the channel capacity, i.e. the maximum \( z_0/\ln 2 \) of the entropy \( H/\ln 2 \), we define

\[ T = \{1, \ldots, d\} \cup \{d+1, \ldots, k+1\} \]
\[ \cup \{k+2, k+3, \ldots\} =: S_1 \cup S_2 \cup S_3, \]
where \( d = 0, 1, \ldots, k = d+1, d+2, \ldots \) are given, and we compute the capacity for the case that the probabilities \( P_1, P_3 \) assigned to the sets \( S_1, S_3 \) are both small. Clearly, the quantities \( P_1 \) and \( P_3 \) denote the probabilities that phrases are transmitted that are either too short or too long, respectively. We find that the familiar capacities of \((d, k)\)-constrained sequences [2] are approached from above as \( P_1, P_3 \to 0 \) with an error \( A(P_3 \ln P_1 + P_3 \ln P_3) \), where we can evaluate the \( A \) explicitly. We obtain a similar result for the case that \( T \) is as in (6) with \( S_1, S_3 \) merged into a single set \( S_1 \cup S_3 \). Further results are published in [3].

Conclusions

We have presented an analytic expression for the maximum of the normalized entropy \( -\sum_{i\in T} p_i \ln p_i / \sum_{i\in T} p_i \), under the condition that \( T \) is the disjoint union of sets \( S_n \) of positive integers that are assigned probabilities \( P_n, \sum_n P_n = 1 \). We computed the capacity of weakly \((d, k)\)-constrained sequences that are allowed to violate the \((d, k)\)-constraint with given probability.

References