MIT/LCS/TM-212

MIMIMAX OPTIMAL UNIVERSAL CODEWORD SETS

Peter Elias

January 1982

MINIMAX OPTIMAL UNIVERSAL CODEWORD SETS

PETER ELIAS
Department of Electrical Engineering and Computer Science
and
Laboratory for Computer Science
Massachusetts Institute of Technology

Abstract

In an interactive multi-user data-processing system a user knows the probabilities of his messages and must encode them into a fixed system-wide variable-length codeword set. He needs to receive the answer to his last message before selecting the next, so his encoding is one-shot. To minimize average codeword length he encodes his messages in order of decreasing probability into codewords in order of increasing length. I give an algorithm which, for each of several measures of performance, finds the codeword set best by that measure for the worst user, and some of the minimax optimal codeword sets the algorithm has found. Some of the results hold for all user distributions: others require e.g. that all users send exactly or at most m distinct messages, and/or that there is an integer k such that no user has a message of probability greater than 1/k.

Introduction.

In an interactive multi-user data-processing system each user or user group may have a different message set or probability distribution, but it may be convenient for the system to require that each user encode his messages into a fixed system-wide set of codewords. Since a user may need to receive the answer to his last message before sending the next, his encoding must be one-shot. I discuss in this paper the problem of finding the best codeword set (in a minimax sense) to use in such a system. For convenience in summarizing previous work [1—4] and to establish notation I first give some known one-user variable-length coding results and define the multi-user problem.

1. Some known one-user variable-length coding results.

Let $p(j)$ denote the probability of the jth most probable message in a countable set M of m messages with $p(j) \geq 0$ for $j < m+1$ and (if $m < \infty$) $p(j) = 0$ for $j > m$. Let U_m denote the class of all such nonincreasing distributions on N^+ which vanish on integers $> m$. Then $p \in U_m$ and $\|p\| \leq m$, where

$$\|p\| = |\{\, j \in N^+ \mid p(j) > 0 \,\}| \text{ for } p \in U. \tag{1.1}$$

Let V_m be the subset of U_m with $p(1) < 1$ (i.e. with $\|p\| \geq 2$) and let W_m be the subset of U_m with $\|p\| = m$. Let $U = U_\infty$, $V = V_\infty$, $W = W_\infty$.

Let C be a prefix-free set of m codewords in a finite alphabet B of size $|B| = b$, $2 \leq b < \infty$, whose jth shortest codeword c_j has length $|c_j| = \lambda(j)$. There is known to be such a set iff the length function λ of

[1]This work was supported in part by the Army Research Office under contract DAAG29 77 C 0012.

C satisfies the Kraft inequality

$$\sum_{1}^{m} b^{-\lambda(j)} \leq 1. \tag{1.2}$$

Let Λ_m denote the class of nondecreasing integer-valued functions on the positive integers $< m+1$ which satisfy (1.2), and $\Lambda = \Lambda_\infty$.

Encoding the m messages in M with distribution $p \in U_m$ in order of decreasing probability $1-1$ into the m codewords in C with length function $\lambda \in \Lambda_m$ in order of increasing length gives an expected cost of

$$E_p(\lambda) = \sum p(j)\lambda(j) \tag{1.3}$$

b-ary symbols per message, and no $1-1$ mapping of M into C has smaller average cost. Doing $1-1$ encoding only for those messages with positive probabilities is possible iff $\|p\| \leq m$ and then has an expected cost also given by (1.3). (In (1.3) and throughout expectations are summed over all and only the nonvanishing values of $p(j)$.)

For $p \in U$ let

$$\begin{aligned} h_m(p) &= \inf_{\lambda \in \Lambda_m} E_p(\lambda) \text{ for } m \geq \|p\|; \\ h(p) &= h_{\|p\|}(p). \end{aligned} \tag{1.4}$$

In the standard one-shot variable-length source encoding problem a distribution $p \in U$ is given and the code designer has the task of constructing a length function λ which approaches or attains the infimum $h(p)$ in (1.4). If $\|p\|$ is finite so is $h(p)$ and the infimum is attained by the length function λ of a Huffman code for p. If $\|p\|$ is infinite and $h(p)$ is finite Huffman coding may still be possible (Humblet [5]), but in any case for every $\epsilon > 0$ there is a function $\lambda \in \Lambda_\infty$ which approaches that infimum, with $E_p(\lambda) < h(p) + \epsilon$. Therefore $h(p)$ is an appropriate measure of the expected transmission or storage cost of an optimal one-shot encoding of messages from M with distribution p into codewords in B^*.

In the standard block-to-variable-length encoding problem the designer maps sequences of k messages from M into a set of m^k codewords in B^*. If successive messages are selected independently from p an optimal encoding takes an average number of symbols per message which approaches the entropy $H(p)$ as k increases, where

$$H(p) = -\sum p(j) \log p(j) = E_p(-\log p). \tag{1.5}$$

Therefore $H(p)$ is an appropriate measure of the cost of an optimal block encoding of sequences of messages from M into codewords in B^* when only the first-order distribution p is given. (In (1.5) and throughout all logarithms are taken to the base $|B| = b \geq 2$.)

There is no explicit formula for $h(p)$ but the values of h and the entropy function H are related. H is bounded by

$$\log 1/p(1) \leq H(p) \leq \log\|p\| \text{ on } U \tag{1.6}$$

and the inequalities are both strict unless $p = q_m$, the uniform distribution on the first m integers, where

$$q_m(j) = \begin{cases} 1/m, & j \leq m < \infty; \\ 0, & m < j, \end{cases} \tag{1.7}$$

in which case there is equality in both inequalities. And $h(p)$ is bounded on U_m by

$$H(p) \leq h(p) \leq h_m(p) < 1 + H(p). \tag{1.8}$$

There is equality in the leftmost inequality iff $p = q_m$ for $m = b^a$, an integral power of b, including the case $a = 0, m = 1, H(q_1) = h(q_1) = 0$. As $\epsilon \to 0$ the distribution p with $p(1) = 1 - \epsilon, p(2) = \epsilon, H(p) \to 0$ and $h(p) = 1$ shows that the constant 1 on the right cannot be improved.

Proposition 1 summarizes some of these results for reference.

Proposition 1.

(i) For every $p \in U_m$, $n \geq m$ *and* $\lambda \in \Lambda_n$, $E_p(\lambda) \geq h_n(p) \geq h_m(p) \geq h(p) \geq H(p)$.

(ii) For every $p \in W_m$ *and* $\epsilon > 0$ *there exists a* $\lambda \in \Lambda_m$ *with* $E_p(\lambda) < 1 + H(p), \quad E_p(\lambda) < \epsilon + h(p)$.

(iii) For every $\epsilon > 0$ *there exists a* $p \in U$ *with* $h(p) > (1 - \epsilon) + H(p)$.

2. Multi-user variable-length coding: universality and cost measures.

Consider next a multi-user version of the one-shot variable-length coding problem in which there is a set of actual or possible users characterized by a class $S \subseteq U$ of nondecreasing distributions. Let

$$s = \|S\| = \sup_{p \in S} \|p\| \text{ for } S \subseteq U. \tag{2.1}$$

The designer constructs as before a single set C of codewords with length function $\lambda \in \Lambda_s$. A user with distribution $p \in S$ assigns his messages in order of decreasing probability to codewords in C in order of increasing length at an expected codeword length of $E_p(\lambda)$ symbols per message. The designer's task is to find a λ whose cost by some measure is reasonable for all users rather than optimal for one of them. This task is closely related to the design of "a universal code" for a class S of nonincreasing distributions, but in the usage which has become standard "a universal code" denotes a family of codes which encode successively larger segments of source output at successively smaller costs per message while in this case the encoding is constrained to be one-shot. ([7] and [8] give principal results and extensive references on universal coding.)

A number of plausible cost measures for the multi-user problem have been explored in part in [1], [2], [3] and [4]. $E_p(\lambda)$ itself is an absolute cost measure. The difference $E_p(\lambda) - h(p)$ and the ratio $E_p(\lambda)/h(p)$ compare (in different ways) the cost of one-shot encoding into C to the cost of a one-shot encoding optimal

for p alone. The difference $E_p(\lambda) - H(p)$ and the ratio $E_p(\lambda)/H(p)$ compare the cost of using C to the cost of an optimal block encoding for a sequence of independent selections from p.

More generally, define the codeword set C with length function $\lambda \in \Lambda_s$ to be *universal* on S iff there are pairs (d, r), (D, R) of nonnegative real constants such that

$$E_p(\lambda) \leq d + rh(p) \text{ and} \tag{2.2a}$$
$$E_p(\lambda) \leq D + RH(p) \text{ for all } p \in S. \tag{2.2b}$$

Then d and D bound the absolute cost measure $E_p(\lambda)$ on S when $r = R = 0$, and bound the two difference measures when $r = R = 1$, while r and R bound the two ratio measures when $d = D = 0$. This definition requires that $E_p(\lambda)$ be finite on S only when $h(p)$ and $H(p)$ are. By the inequalities (1.8), (2.2a) holds for some finite (d, r) iff (2.2b) holds for some finite (D, R), so the definition is unambiguous. Proposition 2 shows that it is not vacuous: i.e. that there are codeword sets universal on U itself, and therefore on any subset S of U.

Proposition 2.

(i) If (2.2) holds at $S = U$ then $D \geq 1$, $d \geq 1$, $r > 1$ and $R > 1$.

(ii) For every $b \geq 2$ there is a $\lambda \in \Lambda$ such that (2.2) holds with $S = U, D = d = 1$ and $R = r = 2$.

Comment. The fact that $D \geq 1$ and $d \geq 1$ follows since $h(q_1) = H(q_1) = 0$ (as noted after (1.8)), while by (1.2) $E_p(\lambda) \geq E_{q_1}(\lambda) = \lambda(1) \geq 1$ for $\lambda \in \Lambda$ (or indeed for $\lambda \in \lambda_m$ and $m > 1$). Davisson and Leon-Garcia prove that $R > 1$ in [4], and $r > 1$ follows from (1.8) for $p = q_n$ and $n \to \infty$. (ii) is proved by construction of an appropriate countably infinite codeword set in [2]. The values $R = r = 2$ are not best possible: see Proposition 5 below. []

Given the existence of codeword sets universal on any $S \subseteq U$ and no *a priori* information about the distribution of users over S, it is natural to look for minimax solutions. When (2.2) holds the cost pair (d, r) is said to be h-attainable and (D, R) to be H-attainable on S, and λ and C are said to attain those costs.

There is no single cheapest attainable cost pair in general, since there are trading relations between the additive and the multiplicative constants in (2.2). The designer may choose to minimize d or D given r or R, or conversely. Let

$$\begin{aligned} d^*(r, S) &= \inf_{\lambda \in \Lambda_s} \sup_{p \in S} \; (E_p(\lambda) - rh(p)), \\ r^*(d, S) &= \inf_{\lambda \in \Lambda_s} \sup_{p \in S} \; (E_p(\lambda) - d)/h(p), \\ D^*(R, S) &= \inf_{\lambda \in \Lambda_s} \sup_{p \in S} \; (E_p(\lambda) - RH(p)), \\ R^*(D, S) &= \inf_{\lambda \in \Lambda_s} \sup_{p \in S} \; (E_p(\lambda) - D)/H(p). \end{aligned} \tag{2.3}$$

Then an h-attainable pair (d, r) and a λ and C which attain it are said to be (minimax) h-optimal on S if $d = d^*(r, S)$ and $r = r^*(d, S)$, and similarly for H-optimality. The minimax value of the absolute cost measure mentioned above is $d^*(0, S) = D^*(0, S)$; the minimax difference measures are $d^*(1, S)$ and $D^*(1, S)$ and the minimax ratio measures are $r^*(0, S)$ and $R^*(0, S)$. Proposition 2 (i) shows that (unlike $r^*(1, S)$ and $R^*(1, S)$ in proposition 2(ii)) these five cost measures all diverge at $S = U$, so it is necessary to consider subsets of U to obtain useful results for them.

The absolute measure and the two difference measures converge on S iff $H(p)$ is finite on S, which holds by (1.6) if s is finite; the ratio measure $r^*(0, S)$ converges iff $h(p) > 0$ on S, which holds on $U - \{ q_1 \}$, and the ratio measure $R^*(0, S)$ converges iff $H(p)$ is bounded below by a constant on S, which by (1.6) holds if $p(1) \leq 1/k$ for some integer $k > 1$ on S. The h-costs are significantly reduced if all users have the same number of messages — i.e. if $\|p\| = s$ on S. Following Rissanen in part I therefore seek optimal cost-pairs and codeword sets on the following subsets of U.

$$\begin{aligned} U_{k,m} &= \{p \in U | 1/k \geq p(1), \|p\| \leq m\}, \qquad 1 \leq k < \infty, \\ V_{k,m} &= \{p \in U | 1/k > p(1), \|p\| \leq m\}, \qquad 1 \leq k < \infty, \\ W_{k,m} &= \{p \in U | 1/k > p(1), \|p\| = m\}, \qquad 1 \leq k < \infty. \end{aligned} \tag{2.4}$$

To be consistent with earlier notation and simplify typography let S_m denote $S_{1,m}$ and let S denote $S_\infty = S_{1,\infty}$, where S is U, V or W. The sets $V_{k,m}$ and $W_{k,m}$ are empty for $k \geq m$. For $m < \infty$ the set $U_{m,m} = \{ q_m \}$, where q_m is the uniform distribution on the first m integers in (1.7), and $V_{1,m} = V_m = U_m - \{ q_1 \}$ so $V = U - \{ q_1 \}$. Note that these definitions extend to $k = 0$ but introduce no new sets: $V_{0,m} = U_{0,m} = U_{1,m} = U_m$, $W_{0,1} = U_1$, $W_{0,m} = W_{1,m}$ for $m \geq 2$. Allowing the value $k = 0$ simplifies the statement of Theorem 2 below.

3. History and literature.

In [1] I constructed a codeword set and proved it to be universal on U in the present sense, with H-costs (2,2) for any $b \geq 2$. I improved that result in [2] to H-costs (1,2), which by (1.9) imply h-costs (1,2), proving Proposition 2(ii). The fact that $h(p) \geq 1$ on $V = U - \{ q_1 \}$ was used to prove that the same set attained h-costs (0,3) on V. The proof in [2] that several codeword sets constructed there (similar to sets constructed earlier by Levenstein [6]) are asymptotically optimal showed that $r^*(0, U_{k,\infty}) \to 1$ as $k \to \infty$. However the techniques used in [2] were not sufficiently precise to determine minimax optimal λ and (d, r) on V or on $U_{k,\infty}$ or to find $r^*(0, V)$ or $r^*(0, U_{k,\infty})$ exactly, or to give smaller cost bounds for finite m.

In [3] Rissanen considered the behavior of the minimax ratio H-cost measure $R^*(0, U_{k,m})$. For $b = 2$, finite m and $k \geq 2$ he showed that $\dot{E}_p(\lambda)/H(p)$ attained its maximum over $U_{k,m}$ on the set $\{ q_k, q_{k+1}, \ldots, q_m \}$ of uniform distributions. Using that result he found an algorithm which gives a lower bound to $R^*(0, U_{k,m})$ and a particular distribution in $U_{k,m}$ whose Huffman encoding has an expected cost which approximates $R^*(0, U_{k,m})$. He obtained lower bounds to $R^*(0, U_{k,m})$ considerably smaller than the upper bound 3 to the value of $r^*(0, V)$ found in [2], for the parameter values $b = 2$, $m = 32$ and $k = 2, 3, 4$ and 6. He conjectured that the difference was due to his finite m. In fact in the binary case, as we shall see below, while $r^*(0, V_{1,m}) \leq R^*(0, V_{2,m}) = R^*(0, U_{2,m})$, the two are very nearly equal and grow very little as m increases from 32 to ∞: the difference between my upper bound and Rissanen's lower bound at $k = 2$

is primarily due to the fact that his bound is closer to the correct value.

In [4] Davisson and Leon-Garcia considered the minimax redundancy H-cost measure $D^*(0, S)$ on a set S of not necessarily monotone distributions. They showed that if S is the convex hull of a set Q of distributions then the difference $E_p(\lambda) - H(p)$ attains its maximum over S on Q. They showed in particular that the difference attains its maximum over the set U_m on the set $\{q_1, q_2, \ldots, q_m\}$, explored the behavior of $D^*(0, U_m)$ for finite m, found an algorithm which gives a lower bound to its value and found a particular distribution in that class whose Huffman encoding has an expected cost which approximates $D^*(0, U_m)$. They showed that $D^*(0, U_m)$ increases without bound as m increases but grows only like $\log \log m$.

The principle results of the present paper are three. Section 4 shows (Theorems 1 and 2) that each of the four suprema over S in (2.3) is the supremum over an appropriate subset of the set $\{q_1, q_2, \ldots, q_s\}$ of uniform distributions when S is any one of the distributions in (2.4), and (Lemma 5) that for h-costs that subset has only $O(\log s)$ members. Those results make it possible to prove the minimax optimality of a number of length funtions. Section 5 derives an algorithm which is shown (Theorem 3) to find cost pairs and length functions in Λ_s which are (h- or H-) optimal, or within any $\epsilon > 0$ of optimal, on the sets $S_{k,m}$ in (2.4). For large m the algorithm converges to exact values much faster than the algorithms given in [3] and [4] converge to approximate values. It finds optimal integer-valued length functions without a separate Huffman coding step and covers h-costs as well as H-costs. Section 6 gives a number of optimal cost-pairs and length functions found by the new algorithm, for finite and infinite m. The Appendix specializes the algorithm for the four functions in (2.3).

4. Extremal properties of uniform distributions.

Theorem 1 gives results for H-costs and Theorem 2 gives similar results for h-costs. These results are the key to the algorithm in Section 5. They also simplify the task of checking how good a proposed length function is in the minimax sense.

Theorem 1.

Let k be an integer and $1 \leq k \leq m \leq \infty$. Let $\lambda \in \Lambda_m$. Let $S_{k,m}$ be any one of $U_{k,m}$, $V_{k,m}$ or $W_{k,m}$ for $k < m$: if $k = m$ let $S_{m,m} = \{q_m\} = U_{m,m}$. Let D and R be real and ≥ 0. Then

$$\begin{aligned} &E_p(\lambda) \leq D + RH(p) \text{ for all } p \in S_{k,m} \text{ iff} \\ &E_{q_j}(\lambda) \leq D + RH(q_j) \text{ for } k \leq j < m + 1. \end{aligned} \tag{4.1}$$

Comment. The case $D = 0, k \geq 2$ is proved in [3], and the case $R = 0, k = 1$ follows from a result in [4]. The proof is new and parallels the proof below of the new results on h-costs. Note that $D \geq D^*(R, S_{k,m})$ and $R \geq R^*(D, S_{k,m})$ iff (4.1) holds. Thus D^* and R^* can be evaluated by finding the smallest value of D (or R) for which the second line in (4.1) holds when the value of R (or D) is given. This is the basis of the algorithms given below. []

Two obvious lemmas are useful in proving Theorem 1.

Lemma 1 (Shannon).

Each $p \in U_m$ is an average of the uniform distributions q_j in (1.7): i.e.

$$p(i) = \sum_{j=1}^{m} w_j q_j(i), \text{ where}$$
$$w_j = j(p(j) - p(j+1)), \tag{4.2}$$

and w is a probability distribution with $\|w\| = \|p\|$.

Proof. Obvious (see Shannon, [9]). []

Lemma 2.

For $p \in U_{k,m}$,

$$E_{q_j}(-\log p) \geq \max\{\log k, \log j\}. \tag{4.3}$$

Proof. $E_{q_j}(-\log p) \geq -\log E_{q_j}(p)$ since $-\log$ is convex $\bigcup$. $E_{q_j}(p) \leq 1/k$ since each $p(j)$ is, and $E_{q_j}(p) = \sum_{i=1}^{j} p(i)/j \leq 1/j$ since all the $p(i)$ sum to 1. []

Proof of Theorem 1.

Let $k \leq j < m+1$. Then $q_j \in U_{k,m}$, so when $S_{k,m} = U_{k,m}$ the first line in (4.1) implies the second. If $S_{k,m}$ is $V_{k,m}$ or $W_{k,m}$ then q_j is not in that set for all those values of j but for any ϵ with $0 < \epsilon < 1$ the distribution q'_j is, where

$$q'_j = (1-\epsilon)q_j + \epsilon p \tag{4.4}$$

and p is any distribution in $W_{k,m}$ with finite entropy (and therefore finite $E_p(\lambda)$ by the first line in (4.1)) — for example $p = q_m$ if $m < \infty$, $p(j) = (b-1)b^{-j}$ if $m = \infty$. By the linearity of the expectation operator E_p in p, the continuity of H and the well-known inequalities

$$\begin{aligned}(1-\epsilon)H(p) + \epsilon H(q) &\\ \leq H((1-\epsilon)p + \epsilon q) &\\ \leq (1-\epsilon)H(p) + \epsilon H(q) + \epsilon\log(1/\epsilon) + (1-\epsilon)\log(1/(1-\epsilon)),&\end{aligned} \tag{4.5}$$

$E_{q'_j}(\lambda) \to E_{q_j}(\lambda)$ and $H(q'_j) \to H(q_j)$ as $\epsilon \to 0$, so the first line in (4.1) implies the second again.

To complete the proof assume that the first line in (4.1) does not hold. Then for some $p \in S_m$ there is an $\delta > 0$ with

$$\begin{aligned}D + \delta < E_p(\lambda) - RH(p) &= E_p(\lambda) - RE_p(-\log p)\\ &= \sum_j w_j(E_{q_j}(\lambda) - RE_{q_j}(-\log p))\\ &\leq \sup_{1 \leq j < m+1} (E_{q_j}(\lambda) - R\max\{\log j, \log k\})\end{aligned} \tag{4.6}$$

using Lemmas 1 and 2 and the fact that an average of terms is no larger than the largest term.

But $E_{q_j}(\lambda)$ increases with j if λ does, so the sup cannot occur for $j < k$. Therefore $j \geq k$ and the max is $\log j = H(q_j)$ so

$$D + \delta \leq \sup_{k \leq j < m+1} (E_{q_j}(\lambda) - RH(q_j)), \tag{4.7}$$

and the second line does not hold, proving the Theorem. []

There is a corresponding theorem for the two other cost measures. It is more complex because the analog to Lemma 2 for h is more complex than Lemma 2. The tightest results hold only on the sets $V_{k,m}$ and $W_{k,m}$ and only when $k = 0$ or k has the special form $k = b^a - (b-1)$ for some $a \in N^+$. Fortunately this includes the important special cases V_m and W_m at $a = k = 1$, and $U_m = V_{0,m}$ (as noted following (2.4)) at $a = k = 0$.

Theorem 2.

Let $a = k = 0$ or let $a \in N^+$ and $k = b^a - (b-1) < m \leq \infty$. Then

(i) If $m \leq b^a$ then the constant length function $\lambda(j) = a$ is in Λ_m and is h-optimal on U_m, V_m and W_m and their nonempty subsets in (2.4), with h-costs $(0, 1)$: i.e.

$$E_p(\lambda) = a = h(p) \text{ on } U_m.$$

(ii) If $b^a < m$ let $\lambda \in \Lambda_m$. Then

$$\begin{aligned} &E_p(\lambda) \leq d + rh(p) \text{ for all } p \in W_{k,m} \text{ iff} \\ &E_{q_j}(\lambda) \leq d + rh_m(q_j) \text{ for } \max\{1, b^a - 1\} \leq j < m+1; \\ &E_p(\lambda) \leq d + rh(p) \text{ for all } p \in V_{k,m} \text{ iff} \\ &E_{q_j}(\lambda) \leq d + rh(q_j) \text{ for } b^a \leq j < m+1. \end{aligned} \tag{4.8}$$

Comment. Since $h(q_j) \geq H(q_j)$, a comparison of the last line of (4.1) with $k = b^a$ to the last line of (4.8) shows that $E_p(\lambda) \leq d + rh(p)$ on $V_{b^a-(b-1),m}$ if $E_p(\lambda) \leq d + rH(p)$ on $U_{b^a,m}$, so that in particular at $a = k = 1$ and $b = 2$, $d^*(r, V_m) \leq D^*(r, U_{2,m})$ and $r^*(d, V_m) \leq R^*(d, U_{2,m})$. Note also that taking $k = a = 0$ in the third and fourth line gives conditions that λ has h-costs $\leq (d, r)$ on $V_{0,m} = U_m$. []

The proof of Theorem 2(ii) requires a lemma analagous to Lemma 2, with the expectation of $-\log p$ replaced by the expectation of the length function μ of an optimal code for p.

Lemma 3.

Let p be in $V_{k,m}$, $m > k \geq 0$ and let $a = \lfloor \log(k + b - 1) \rfloor$, a nonnegative integer. Then $p(1) < 1/k$: choose $\epsilon \geq 0$, < 1 such that $\epsilon < 1/k - p(1)$. Let $\mu \in \Lambda_m$ be optimal or nearly optimal for p, satisfying $E_p(\mu) < h_m(p) + \epsilon$. Then

$$E_{q_j}(\mu) \geq max\{\, a, h_m(q_j)\,\},$$
$$E_p(\mu) \geq a. \tag{4.9}$$

Proof of Lemma 3. If $k = 0$ then $a = 0$ and (4.9) holds trivially, by the nonnegativity of μ and h_m and the fact that μ is no better than optimal for q_j. Thus assume $k > 0, a > 0$. The existence of μ given ϵ follows from Proposition 1(ii) above.

In the b-ary tree of a prefix-free codeword set for p with increasing length function μ define the set S of all leaves at level $\mu(1)$ and all (leaf or interior) vertices at level $\mu(1) + 1$ with no ancestor in S. Label each vertex in S with the sum of the probabilities of the codewords for which that vertex is a prefix. The labels sum to 1 since S is a cutset. Let $V1$ denote the leaf for the codeword c_1 at level $\mu(1)$ with the largest probability $p(1)$, and let N_1 denote the total number of leaves at level $\mu(1)$.

No vertex V_2 in S at level $\mu(1) + 1$ has probability $> p(1) + \epsilon$ since if it did, using V_2 for c_1 and moving the subtree rooted at V_2 to a root at V_1 would leave a code with length function μ' and $E_p(\mu') < E_p(\mu) - \epsilon < h_m(p)$, which violates Proposition 1(i). There are N_1 leaves at level $\mu(1)$, with probabilities at most $p(1)$, so there are at most $(b^{\mu(1)} - N_1)b$ vertices at level $\mu(1) + 1$ in S with total probability at least $1 - N_1 p(1)$. Therefore

$$\begin{aligned} (p(1) + \epsilon)(b^{\mu(1)} - N_1)b \geq 1 - N_1 p(1) > 1 - N_1(p(1) + \epsilon), \text{ so} \\ (p(1) + \epsilon) > 1/(b^{\mu(1)+1} - N_1(b - 1)), \text{ but} \\ 1/(b^a - (b - 1)) \geq 1/k \geq p(1) + \epsilon \end{aligned} \tag{4.10}$$

and $N_1 \geq 1$ so $\mu(1) \geq a$. Since μ is nondecreasing, $E_p(\mu) \geq a$ and $E_{q_j}(\mu) \geq a$. Since μ is no better than optimal in Λ_m for q_j, $E_{q_j}(\mu) \geq h_m(q_j)$. []

One additional Lemma is required in proving Theorem 2. A simple tree construction gives the values of the sum of the codeword lengths in a codeword set in Λ_m or in Λ_j which is optimal for the uniform distribution q_j in (1.7).

Lemma 4.

Let $L(j) = \lfloor \log j \rfloor$. *Then*

$$\begin{aligned} jh(q_j) = jh_j(q_j) = \lceil (j - b^{L(j)})(L(j) + b/(b-1)) \rceil + L(j)b^{L(j)}, \qquad b^{L(j)} \leq j < b^{L(j)+1}; \\ jh_m(q_j) = \left\{ \begin{matrix} jh(q_j) + 1 \text{ if } j \equiv 1 \pmod{b-1}, \\ jh(q_j) \text{ otherwise} \end{matrix} \right\}, \qquad 1 \leq j < m. \end{aligned} \tag{4.11}$$

Proof of Theorem 2

(i) The given λ clearly attains $E_p(\lambda) = a$ on $V_{0,m} = U_m$, and by Lemma 3 that result cannot be improved on any subset of U_m.

(ii) By hypothesis $m \geq 1$. Assume first that p is in $W_{k,m}$ and let μ be a length function in Λ_m which is approximately optimal for p so that $E_p(\mu) - \epsilon \leq h(p) = h_m(p)$. By Lemma 1 express $E_p(\lambda)$ and $E_p(\mu)$ as averages of the expectations of λ and μ with respect to the q_j. Then a derivation which parallels (4.6) using Lemma 3 rather than Lemma 2 shows that if the first line in (i) does not hold then for some $\delta > 0$

$$\begin{aligned} d + \delta < E_p(\lambda) - rh(p) &\leq E_p(\lambda) - r(E_p(\mu) - \epsilon) \\ &= \sum_j w_j(E_{q_j}(\lambda) - r(E_{q_j}(\mu) - \epsilon)) \\ &\leq \sup_{1 \leq j < m+1} (E_{q_j}(\lambda) - r(\max\{a, h_m(q_j)\} - \epsilon)). \end{aligned} \tag{4.12}$$

For $k = 0$, $a = 0$ and the max in (4.12) is $h_m(q_j)$. For $k > 0$, $a > 0$ and $h_m(q_j)$ is nondecreasing in j and by (4.10) is $\leq a$ for $j \leq b^a - 1$, which is $< m$. Therefore in (4.12) the max is a for $j \leq b^a - 1$, while the expectation of λ increases with j so that the sup of the difference occurs at $j \geq \max\{1, b^a - 1\}$ and

$$d + \delta - r\epsilon \leq \sup_{\max\{1, b^a - 1\} \leq j < m+1} (E_{q_j}(\lambda) - rh_m(q_j)). \tag{4.13}$$

Taking $r\epsilon < \delta$ shows that the second line in (ii) does not hold, so it implies the first line, proving half of (ii) for $W_{k,m}$. The converse implication follows from the fact that q_m is in $W_{k,m}$ while for $j < m$ the q' in (4.4) is in $W_{k,m}$, as before, which completes the proof of the equivalence of the first two lines in (4.8).

If p is in $V_{k,m}$ its first t terms are a distribution in $W_{k,t}$ for some $t < m + 1$ so if the third line in (ii) does not hold then (4.12) holds with m replaced by t and $h_m(q_l)$ replaced by $h_l(q_l) = h(q_l)$ for all $k < t < m + 1$. Since $h(q_t) \leq h_m(q_t)$ changing h_m to h for all $j \geq 1$ strengthens the inequalities, so if the third line fails then for small ϵ (4.12) gives

$$d < \sup_{1 \leq j < m+1} (E_{q_j}(\lambda) - r \max\{a, h(q_j)\}). \tag{< 4.14}$$

But from (4.11), $h(q_j) \leq a$ for $j \leq b^a$, so the max is constant while the expectation increases with j and the sup cannot occur before $j = b^a$, which gives

$$d < \sup_{b^a \leq j < m+1} (E_{q_j}(\lambda) - rh(q_j)), \tag{4.15}$$

and shows that the fourth line in (ii) implies the third. The converse implication holds since q_j is in $V_{k,m}$ for $b^a \leq j < m + 1$, which proves the equivalence of the last two lines in (4.8). []

A final Lemma produces a dramatic reduction in the number of uniform distributions which must be considered in evaluating r^* and d^* on $V_{k,m}$ and $W_{k,m}$ for large m.

Lemma 5.

Let $a \in N$ and $m > b^a$. Let $L(j) = \lfloor \log j \rfloor$. Then

(i) The inequality $E_{q_j}(\lambda) \le d + rh(q_j)$ *holds for all* j *with* $b^a \le j < m+1$ *iff it holds for all* $j < m$ *with* $j = b^s$ *for integer* $s \ge a$ *and, if* $m < \infty$, *at* $j = m_0$ *and at* $j = m$, *where* $m_0 = m - \rho$ *and* ρ *is the remainder of* $m - 1 \pmod{b-1}$ *(and thus is* 0 *if* $b = 2$ *or if* m *is an integral power of* b*).*

(ii) Let $b = 2$. *Then the inequality* $E_{q_j}(\lambda) \le d + rh_m(q_j)$ *holds for all* j *with* $2^a - 1 \le j < m+1$ *iff it holds for* $j = 2^a - 1$, *for all* $j < m$ *with* $j = 2^s$ *for some integer* $s \ge a$ *and, if* $m < \infty$, *at* $j = m_0$ *and at* $j = m$, *where* $m_0 = b^{L(m-1)}$.

(iii) Let $b > 2$. *Then the inequality* $E_{q_j}(\lambda) \le d + rh_m(q_j)$ *holds for all* j *with* $b^a - 1 \le j < m+1$ *iff it holds for all* $j < m_0$ *with* $j = b^s - 1$ *or* $j = b^s + b - 2$ *for some integer* $s \ge a$ *and, if* $m < \infty$, *at* $j = m_0$ *and at* $j = m$, *where now* ρ *is the remainder of* $m \pmod{b-1}$ *and*

$$m_0 = \begin{Bmatrix} b^{L(m-1)} + b - 2 \text{ if } \rho \le 1, \\ m - \rho \text{ otherwise} \end{Bmatrix}.$$

Proof. (i) Multiply both sides of the inequality by j and observe that in Lemma 4 dropping the ceiling corners in (4.11) gives a piecewise linear lower bound to $F(j) = jh(q_j)$, of slope $f(j) = F(j) - F(j-1) = L(j-1) + b/(b-1)$. $F(j)$ is equal to this lower bound when $j \equiv 1 \pmod{b-1}$ (and therefore for all j when $b = 2$, and for all $j = b^s$ for any b). If $m < \infty$ and $\rho > 0$ then another bounding linear segment goes from $m - \rho$ to m. The term jd is also linear in j. But $jE_{q_j}(\lambda)$ is just the sum of the first j values of the nondecreasing function λ, so it is convex $\bigcup$. It follows that that sum lies below $jd + rjh(q_j)$ at intermediate j if it does so at two points which lie on the same linear piece of the lower bound and attain equality in that bound, as do b^s and b^{s+1}, and (when $m < \infty$) $b^{L(m-1)}$ and $m - \rho$, and (when $\rho > 0$) $m - \rho$ and m.

The same argument holds for (ii) and (iii), but the piecewise linear lower bound to the right side of the inequality changes.

(ii) For $b = 2$, $jh_m(q_j) = jh(q_j) + 1$ for all $j < m$. The argument $2^a - 1$ is added to the set in (i) on which the inequality must hold: the bound to $jh_m(q_j)$ has the same slope as in (i) and its value is increased by 1. For $m < \infty$, in the interval from m_0 to m the slope of the linear lower bound is reduced because $jh_m(q_j) = jh(q_j)$ at $j = m$.

(iii) For $b > 2$, $jh_m(q_j) = jh(q_j) + 1$ for $j < m$ only when $j \equiv 1 \pmod{b-1}$. A careful inspection of (4.11) shows that $F(j) = jh_m(q_j)$ has two segments of linear lower bound per octave, of slopes $f(j) = F(j) - F(j-1) = s+1$ from $j = b^s$ to $j = b^s + b - 2$ and $f(j) = s + b/(b-1)$ from $j = b^s + b - 1$ to $b^{s+1} - 1$. Those bounds are attained when $j \equiv 0 \pmod{b-1}$ (and thus at all the arguments specified in (iii) above except perhaps m). When $\rho = 1$ there is a linear lower bound of smaller slope in the last octave as in (ii): if ρ is neither 1 nor 0 then there is a linear bound of larger slope from $m - \rho$ to m as in (i). []

Unfortunately there is no similar result for D^* and R^*, because the function $jH(q_j)$ is strictly convex $\bigcup$ and cannot be lower-bounded by a set of one or two of its secants per octave. As a result some nicely structured h-optimal length functions are given in closed form in Section 6 but there are few such results for H-optimal length functions. However $jH(q_j)$ can be bounded below by a set of its tangents, and the algorithm given next is equally fast in both cases although it suggests fewer closed-form results in the H-optimal case.

5. A fast algorithm for minimax solutions.

Theorems 1 and 2 allow the reduction of the problems of finding minimax values of the various cost measures on the various sets of distributions mentioned there to a common form. As a first step consider

Problem 1.

Find the smallest x (or y) ≥ 0 for a given y (or x) ≥ 0 such that there is an integer-valued nondecreasing function λ on integers $< m+1$ which satisfies the set of inequalities

$$jE_{q_j}(\lambda) = \sum_1^j \lambda(i) \leq xj + yF(j), \qquad k \leq j < m+1, \tag{5.1}$$

and the inequality

$$\sum_1^m b^{-\lambda(j)} = \sigma(\lambda) \leq 1, \tag{5.2}$$

and find a λ which attains that minimum.

Taking $F(j)$ to be $jH(q_j) = j \log j$ or $jh(q_j)$ or $jh_m(q_j)$, x to be d or D and y to be r or R, Theorems 1 and 2 show that the minimizing x's and y's include the values of D^*, R^*, d^* and r^* on the sets $S_{k,m}$ mentioned in those theorems.

The sum on the left in (5.1) is convex $\bigcup$ since λ is nondecreasing. It follows that (5.1) holds iff

$$\begin{aligned} &\sum_1^j \lambda(i) \leq xj + yF'(j) = xj + y\sum_1^j f(i) \text{ for } 1 \leq j < m+1, \text{ where} \\ &F'(j) = \begin{cases} jF(k)/k, & 0 \leq j \leq k, \\ F(j), & k < j < m+1, \end{cases} \\ &f(j) = F'(j) - F'(j-1), \qquad 1 \leq j < m+1. \end{aligned} \tag{5.3}$$

Assume for the present that F is convex $\bigcup$ as is $j \log j$ (and in the binary case $jh(q_j)$). Then f is nondecreasing. Let λ' be permitted to take real values but satisfy the other requirements on the integer-valued λ in Problem 1. Since the Kraft sum $\sigma(\lambda')$ is decreasing and convex $\bigcup$ in the values of λ', if λ' satisfies both (5.3) and (5.1) for a given x and y it will be possible to decrease x or y or $\lambda'(j)$ for some j unless there is equality in all inequalities. Thus a pair (x, y) is minimal and that minimum is attained by a nondecreasing real-valued function λ' iff

$$\lambda'(j) = x + yf(j), \tag{5.4}$$

where the values of x and y are related by

$$b^x = \sum_1^m b^{-yf(j)}. \tag{5.5}$$

Since λ' is less constrained than λ, fixing x or y in (5.5) and finding y or x gives a lower bound to one of the cost measures. The lower bounds in [3] and [4] on $R^*(0, U_{k,m})$ and $D^*(0, U_m)$ are proved there by specializations of this analysis to $F(j) = j \log j$ and $x = 0$ or $y = 1$.

Solving (5.5) for x given y when $F(j) = j \log j$ requires evaluating and summing $m - k$ distinct terms. Solving for y given x requires such an evaluation and summing for each trial y: using a Newton-Raphson approach, $\log n$ such evaluations are required to obtain an n-bit approximation to the lower bound, or $O((m-k)\log n)$ steps in all. This is not an excessive amount of computation for small m, but becomes noticeable in e.g. designing a codeword set to be used in representing all the integers whose standard binary representation takes 36 bits. And the resulting λ' is not directly attainable, while the Huffman encoding of the distribution $p(j) = b^{-\lambda'(j)}$ is not easy to carry out for large m.

The solution of (5.5) is faster in the binary case if $F(j) = jh(q_j)$, since then by Lemma 4 the difference function f takes only one value , $t + 2$, for each octave of arguments, $b^t < j \leq b^t + 1$, so only $\log m$ terms need be summed. However the result still gives only a lower bound to r^* or D^* and does not give a length function directly since the solution is not integer-valued.

An algorithm which solves Problem 1 as stated, to any desired accuracy, is almost as fast as the approximate solution in this special case. That algorithm requires solving a subproblem n times in order to obtain an n-bit value of one of the cost measures d^*, r^*, D^* , R^* on one of the sets $S_{k,m}$ in theorems 1 and 2. Problem 2 is the subproblem.

Problem 2.

Let g be positive and nondecreasing on the positive integers $< m+1$. Find an integer-valued nondecreasing function λ on that domain which minimizes the Kraft sum $\sigma(\lambda)$ in (5.2) subject to the constraints

$$\sum_1^j \lambda(i) \leq \sum_1^j g(i) = G(j) \text{ for } 1 \leq j < m+1, \tag{5.6}$$

and find the value of that minimal sum.

To solve Problem 1 for an F which is convex $\bigcup$ and a given y (or x) one first solves Problem 2 for the function $g = x + yf$ and the sum $\sigma(\lambda)$; using the f given by (5.3), the given y (or x) and a trial value of x (or y). A geometric search then replaces the trial x by an n-bit approximation to the smallest x with $\sigma(\lambda) \leq 1$ in n solutions to Problem 2. (The variation of $\sigma(\lambda)$ with the trial values may not be smooth enough to give the more rapid Newton-Raphson convergence.)

To solve Problem 2 given g, note that the first sum in (5.6) is integer-valued and convex $\bigcup$ and is therefore the supremum of a set of integer-valued linear functions

$$\sum_1^j \lambda(i) = \sup\{\, \alpha_k(j) \mid k \in N \,\}, \qquad 0 \le j < m+1, \tag{5.7}$$

where α_k has integral slope k and $\alpha_k(j) \le G(j)$. Since $\sigma(\lambda)$ in (5.2) is convex $\bigcup$ and decreasing in the values of λ, $\sigma(\lambda)$ is minimized by choosing each α_k to be as large as possible subject to those constraints.

An integer-valued line of slope k lying below the convex $\bigcup$ function G for arguments $< m+1$ comes no closer to G than it does at its "tangent" argument $j = t(k)$, where

$$t(k) = \min\{\, m, \inf\{\, j \in N | k \le g(j+1) \,\}\,\}, \qquad 0 \le k < \infty, \tag{5.8}$$

and the largest such line takes the value $\lfloor G(t(k)) \rfloor$ at that point, so

$$\begin{aligned} \alpha_k(j) &= (j - t(k))k + \lfloor G(t(k)) \rfloor \\ &= jk - s(k), \text{ where} \\ s(k) &= kt(k) - \lfloor G(t(k)) \rfloor. \end{aligned} \tag{5.9}$$

The line α_k lies above α_{k+1} until they cross at an argument $j = j_k$. Setting $\alpha_k(j_k) = \alpha_{k+1}(j_k)$ determines j_k and the value of the distribution function ϕ_λ of λ:

$$\begin{aligned} j_k &= s(k+1) - s(k), \\ \phi_\lambda(k) &= |\{\, j \in N \mid \lambda(j) = k \,\}| = j_k - j_{k-1} \\ &= s(k+1) - 2s(k) + s(k-1). \end{aligned} \tag{5.10}$$

Theorem 3 summarizes the results.

Theorem 3.

Let g satisfy the requirements of problem 2. Then ϕ_λ in (5.10) is the distribution function of the unique integer-valued nondecreasing function λ which solves the problem. The minimum sum in (5.2) is

$$\sigma(\lambda) = \sum_{j=1}^{m} b^{-\lambda(j)} = \sum_{k=0}^{\infty} \phi_\lambda(k) b^{-k}. \tag{5.11}$$

Given any g satisfying the requirements of Problem 2, Theorem 3 solves that problem in $O((g(m) - g(1)) \log m)$ steps. For by (5.8) and (5.9) $t(k) = s(k) = 0$ if $k \le g(1)$, and if $k > g(m)$ then $t(k) = m$ so $s(k)$ is linear in m and the second difference $\phi(k+1) = 0$. Therefore it is only necessary to find the value of $t(k)$ when $g(1) < k \le g(m)$, a total of $< g(m) - g(1)$ values each $\le m$. Since g is nondecreasing, at worst a binary search finds each value of $t(k)$ in $O(\log m)$ steps. The total number of steps required to find an n-bit solution to Problem 1 by binary search is therefore at most $O((g(m) - g(1))n \log m)$.

The computation is even faster for the g functions needed to evaluate the minimax values of the relative performance measures d^*, r^*, D^* and R^* on the sets $S_{k,m}$ in Theorems 1 and 2. For the function $g(i)$ in those cases can be inverted to give the range of i values for which $g(i) < t$, and $t(i)$ can therefore

be evaluated in $O(1)$ rather than $O(\log m)$ steps. And in all these cases $g(m) - g(1) = O(\log m/k)$, so Problem 2 can be solved in $O(\log m/k)$ steps and an n-bit solution to Problem 1 found in $O(n \log m/k)$ steps. For the h-solutions, if the given x or y is rational the solution is also rational, always when m is finite and often when $m = \infty$, in which case an exact solution can be found with finite computation. The details are given in the appendix. A minor extension is required when the relevant F is not convex $\bigcup$. In this more general case the function F' must be defined not by (5.3) but as the largest convex $\bigcup$ function which lies on or below the point $(0, 0)$ and the set of points $(j, F(j))$ for $k \leq j < m + 1$. Then f is defined as the difference of that F' as in (5.3) and $g(j) = x + yf(j)$ as before. When $F(j)$ is $jh(q_j)$ or $jh_m(q_j)$ the function F' is also the convex hull of the smaller set of points which are the intersections of the segments of the piecewise linear lower bounds in Lemma 5, at which those lower bounds are attained, and its values at intermediate arguments can be computed by interpolation between its values at those intersections.

6. Some minimax optimal cost pairs and distribution functions.

It is not feasible to give values of all four of the minimax cost measures d^*, D^*, r^*, R^* in (2.3) on all three of the sets $U_{k,m}$, $V_{k,m}$, $W_{k,m}$ in (2.4) as functions of the real parameters d, D, r, R and the integer parameters b, k, m. I present here some closed form and some numerical results for the four cases in which the argument d, D, r or R is equal either to 1 or 0, mostly for $b = 2$.

6.1 Results for the absolute measure $d^*(0, S) = D^*(0, S)$.

Codeword sets which are minimax optimal for the absolute measure are well-known. Since $\lambda \in \Lambda_m$ is nondecreasing, if $p \in U_m$ then $E_p(\lambda) \leq E_{q_m}(\lambda)$, so q_m is the worst distribution in any subset $S \subseteq U_m$ to which it belongs. Thus

Proposition 3.

If $S_m \subseteq U_m$ and $q_m \in S_m$, $p \in S_m$ then $E_p(\lambda) \leq E_{q_m}(\lambda)$, so

$$d^*(0, S_m) = D^*(0, S_m) = h(q_m). \tag{6.1}$$

It follows that $d^*(0, S) = \infty$ on U, V and W.

The exact value of $h(q_m)$ can be found from Lemma 4. Since $H(q_m) = \log m$ and $H \leq h < H + 1$, $0 \leq h(q_m) - \log_b m < 1$. The difference is always quite small at $b = 2$, but can approach 1 for some values of m when b is large. Let

$$\gamma(b) = \sup_{m \in N+} (h(q_m) - \log_b m). \tag{6.2}$$

For $b = 2$ a calculus maximization shows that $\gamma(2) = 1 - \log_2 e + \log_2 \log_2 e \approx 0.086$. However $\gamma(b) \geq h(q_2) - H(q_2) = 1 - \log_b 2$, which approaches 1 as $b \to \infty$.

6.2 Results for the difference measures $d^*(1, S), D^*(1, S)$

The problem of finding the minimax optimal length distributions which attain integral values of the difference measure $d^*(1, S)$ for various S — i.e. of finding λ such that $E_p(\lambda) - h(p) \leq d$ on S for integral d — has optimal solutions with simple structure for a number of parameter values. I first give optimal costs in the binary case, and then describe the length functions that attain those costs and some length functions for larger alphabets.

Proposition 4.

Let $b = 2$. For integer $d \geq 0$ let $m = m(d) = 2^{2^{d+2}-4}$. Then

$$d^*(1, U_m) = d^*(1, V_{2m}) = d^*(1, W_{4m-1}) = d. \tag{6.3}$$

It follows that $d^*(1, S) = \infty$ on U, V and W. []

Proposition 4 implies that for $b = 2$ there are optimal codeword sets with $d = 0$ on U_1, V_2 and W_3. The first contains a single codeword, the empty string, of length 0. The second contains the two one-bit codewords. The third has three codewords, one of length 1 and two of length 2.

For integral $d > 0$ the structure of the length functions which attain the costs in (6.3) is most easily shown by giving not λ itself but rather its distribution function ϕ_λ defined in (5.9), which gives the number of codewords of each length.

For $d > 0, b = 2$ and integer $a \geq 0$ the distribution function ϕ_λ which is zero at arguments $< d + a$ and which at arguments $d + a, d + a + 1, \ldots$ takes values

$$2^a, 0, 2^a, 2^{a+1}, 2^{a+2}, \ldots, 2^{2^{d+2}+a-5} \tag{6.4}$$

is minimax optimal on $V_{k,m}$ for $m = m(d) = 2^{2^{d+2}+a-4}$ and $k = 2^a - 1$ for integer a. By Theorem 2 and Lemma 5 that result follows from the easy computation that $E_{q_j}(\lambda) = d + s$ for $j = b^s$ and $a \leq s \leq m(d)$. Note that $V_{k,m}$ reduces to U_m when $a = k = 0$ and to V_m when $a = k = 1$, so (6.4) covers the first two cases in (6.3).

On the smaller sets $W_{k,m}$ of distributions which assign strictly positive probabilities to all m messages, for $b = 2$, $a \geq 1$ and positive integer $d = d^*(1, W_{k,m})$, the optimal distribution function is zero at arguments $< d + a$ and at arguments $d + a, d + a + 1, \ldots$ takes values

$$2^a - 1, 1, 2^a, 2^{a+1}, 2^{a+2}, \ldots, 2^{2^{d+2}+a-5}, 2^{2^{d+2}+a-5} + 1, 2^{2^{d+2}-3} - 2 \tag{6.5}$$

and $m = m(d) = (1 + 2^{-(a-1)})2^{2^{d+2}+a-4} - 1$. In particular at $k = a = 1$, $m(d) = 2^{2^{d+2}-2} - 1$, which covers the last case in (6.3). Again the result may be checked by using Lemma 5.

For an arbitrary alphabet size b I give only the distribution function which generalizes (6.4): its values at arguments $d + a, d + a + 1, \ldots$ are

$$b^a, (b-2)b^a, (b-1)^2 b^a, (b-1)^2 b^{a+1}, \ldots, (b-1)^2 b^{t+a-2}, b^{t+a-1} \tag{6.6}$$

and can be shown by Lemma 5 to be optimal on $V_{k,m}$ with costs $(1, d)$ for $d \in N$ if $k = max\{0, b^a - (b - 1)\}$ for some $a \in N$, $a + d > 0$, $t = (b/(b-1))^2(b^d - 1)$ is an integer (as it is whenever $d \equiv 0 \pmod{b-1}$, and thus for all d when $b = 2$) and $m = m(d) = b^{t+a} = b^{a+(b^d+2-b^2)/(b-1)^2}$.

The algorithm in Section 7 finds minimax optimal length functions which attain positive integral (and other) values of $D = D^*(1, S_{k,m})$ for all $k \in N^+$, and D^* takes the same value whether $S_{k,m}$ is $U_{k,m}$, $V_{k,m}$ or $W_{k,m}$, by Theorem 1. The length functions optimal for D^* have less regular structure than those just given for d^*: I give only the distribution function

$$1, 0, 2, 1, 4, 4 \tag{6.7}$$

for a set of 12 binary codewords which attains $D^*(1, U_{12}) = 1$.

For some values of k the performance of the length function optimal for D^* on $U_{k,m}$ can be bounded without finding it explicitly. It follows from (6.2) above and the comment after Theorem 2 that for $a \in N^+$, $k = b^a$ and $k' = b^a - (b-1)$ and for the length function λ defined by (6.6), which is d^*-optimal for $V_{k',m}$,

$$d^*(r, V_{k',m}) \leq D^*(r, U_{k,m}) \leq \sup_{p \in U_{k,m}} (E_p(\lambda) - rH(p)) \leq r\gamma(b) + d^*(r, V_{k',m}). \tag{6.8}$$

(6.8) not only bounds the performance of the length function optimal for D^* but shows that the λ of (6.5), which is not optimal for D^*, satisfies those bounds. By (6.2) the bounding is quite tight at $b = 2$, $\gamma(2) \approx 0.086$. Computation using the new algorithm shows that the upper bound $\gamma(2)$ to the difference $D^*(1, U_{2,m}) - d^*(1, V_{1,m})$ from (6.8) is in fact approached, e.g. at $m = 2^5$, $d^*(1, V_m) = 1$ and at $m = 2^{13}$, $d^*(1, V_m) = 2$, while the lower bound 0 to that difference is attained, e.g. at $m = 2^4$, $d^*(1, V_m) = 1$, at $m = 2^{12}$, $d^*(1, V_m) = 2$ and at $m = 2^{18}$, $d^*(1, V_m) = 5/2$.

6.3 Results for $r^*(1, S), R^*(1, S)$

The following revised version of Proposition 2(ii) gives optimal values of $r^*(d, S)$ and $R^*(D, S)$ for $b = 2$ when S is U, V or W and $d = D = 1$.

Proposition 5.

Let $b = 2$. Then there are length functions λ, λ', λ'', μ, μ' such that

$$E_p(\lambda) \leq 1 + (207/160)h(p) = 1.29375h(p) \text{ on } U,$$
$$E_p(\mu) \leq 1 + 1.3H(p) \text{ on } U;$$
$$E_p(\lambda') \leq 1 + (5/4)h(p) = 1.25h(p) \text{ on } V,$$
$$E_p(\mu') \leq 1 + 1.26H(p) \text{ on } U_{2,\infty};$$
$$E_p(\lambda'') \leq 1 + (967/800)h(p) = 1.20875h(p) \text{ on } W.$$

Comment. The result that $E_p(\mu) \leq 1 + 1.3H(p)$ on U for the best μ for U is remarkably good compared both to the best possible result $E_p(\mu) \leq 1 + (1 + \epsilon)H(p)$ allowed for that μ by Proposition 2(i) and to the result in Proposition 1 (iii) that $E_p(\mu) > (1 - \epsilon) + H(p)$ for some p in U when μ is the best length function

for p alone. At $b = 2$ the coefficients $207/160$, $5/4$ and $967/800$ are best possible and 1.3 and 1.26 almost best possible, but the best possible coefficients are smaller for larger b. []

The length functions μ, μ', λ, and λ'' are neither simple to describe nor illuminating. However the length funtion for the λ' which is h-optimal on V has some structure. The sequence of values of its distribution function $\phi_{\lambda'}$ at arguments in N^+ is given in (6.9). After the first five values each of the bracketed sequences of five terms is multiplied by 16 to get the next such sequence, so the value sequence is quasigeometric.

$$\begin{aligned} &0, 2, 0, 0, 2, \{4, 6, 10, 16, 24\}, \\ &\qquad \{64, 96, 160, 256, 384\}, \\ &\qquad \{1024, 1536, 2560, 4096, 6144\}, \ldots \end{aligned} \tag{6.9}$$

Using the illustrated structure it is easy to calculate that the Kraft sum is $31/32$. The distribution function with intitial values 0, 2, 0, 1, 1 rather than 0, 2, 0, 0, 2 as in (6.9) has a Kraft sum of 1 and has better performance for almost all users, but is no better in the limit of large k for a user with distribution q_k. It is also easy to check that the given $\phi_{\lambda'}$ satisfies the equation $E_{q_j}(\lambda') = 1 + 1.25h(q_j)$ for $j = 2^k$ and all $k > 1$, which proves the minimax optimality of λ' using Theorem 2 and Lemma 5.

6.4 Results for the ratio measures $r^*(0, S)$, $R^*(0, S)$.

In the case of the ratio measures it is easy to see that the cost pair (d, r) has $d \geq 1$ for any codeword set universal on U, and indeed on any set which includes both the degenerate distribution q_1 and some nondegenerate distribution, e.g. q_2, so $r^*(d, U_m) = \infty$ on U_m if $d < 1$ and $m \geq 2$. There are codeword sets h-optimal on V_m and W_m with $d = 0$ for $2 \leq m \leq \infty$, however. I give the results for $m = \infty$ first and then describe some of the behavior for smaller m.

Proposition 6.

Let $b = 2$. Then

$$\begin{aligned} r^*(0, U) &= \infty \\ r^*(0, V) &= 253/160 = 1.58125, \\ r^*(0, W) &= 559/385 = 1.451948... \end{aligned} \tag{6.10}$$

At $m = 2$ the initial value $r^*(0, V_2) = 1$. There is then a plateau, with $r^*(0, V_m) = 3/2$ for $3 \leq m \leq 53$. A codeword set C optimal for V_{53} has a length-distribution function $\phi_\lambda(j)$ with value sequence

$$1, 1, 0, 1, 1, 4, 4, 4, 16, 11, 10, \tag{6.11}$$

Minimax optimal length functions for $3 \leq m < 53$ are obtained by dropping all but the m shortest codewords from C. C itself attains equality in the Kraft inequality (1.1), and when some words are dropped others can be shortened to maintain that equality: however the resultant codeword set, which will be better for some users, will be no better for the worst user than the shortest m words in C: the minimax value of $E_p(\lambda)$ remains at 1.5 on U_m. The ultimate value $r^*(0, V_m) = 253/160$, not much larger than $3/2$, is first attained for $m \approx 2^{17}$. For $m > 2^{17}$ and $r^*(0, V_m) = 253/160$ the initial values of ϕ_λ are

$$1, 1, 0, 1, 1, 3, 1, 8, 8, 8, 30, 25, 41, 98, 78, \ldots \tag{6.12}$$

The values $r^*(0, W_2) = r^*(0, W_3) = 1$ and there is a plateau with $r^*(0, W_m) = 4/3$ for $11 \leq m \leq 29$. The ultimate value $r^*(0, W_m) = 559/385$ is attained for the first time at $m \approx 2^{30}$. An optimal C for W_{29} has a distribution function whose value sequence is

$$1, 0, 1, 2, 3, 4, 6, 12. \tag{6.13}$$

Again dropping all but the shortest m words from C gives a minimax optimal set on W_m for $11 \leq m < 29$.

The measure $R^*(0, S)$ diverges for $S = U_m$, V_m or W_m since $H(p)$ is not bounded away from 0 on those sets. However there is a bounding relation between $r^*(0, V_m)$ and $R^*(0, U_{2,m})$. An argument like that used in (6.8) shows that if λ is r^*-optimal on $V_{k',m}$ and $k = b^a, k' = b^a - (b-1)$ for integer a then

$$\begin{aligned} r^*(d, V_{k',m}) \leq R^*(d, U_{k,m}) &\leq \sup_{p \in U_{k,m}} ((E_p(\lambda) - d)/H(p)) \\ &\leq \sup_{p \in U_{k,m}} (1 + \gamma(b)/H(p)) r^*(d, V_{k',m}) \\ &= (1 + \gamma(b)/a) r^*(d, V_{k',m}) \end{aligned} \tag{6.14}$$

since $\log k = a$ is the minimum of $H(p)$ on $U_{k,m}$. In particular at $b = 2$ and $k = a = 1$, (6.13) gives $R^*(0, U_{2,m})/r^*(0, V_m) < 1.086$ for all m, and numerical results for $m = 2^j, 1 \leq j \leq 100$ give a tighter bound of 1.016 to that ratio for those values of m.

It is not feasible to give results equivalent to Proposition 6 for all values of the alphabet size b. In general the value of r^* decreases as b increases. Table 1 shows that decrease for $r^*(0, V)$: there is a firm lower bound $r^*(0, V) \geq 4/3$ which is approached as $b \to \infty$: the proof of the bound is left as an excercise for the reader.

TABLE 1

base b	$r^*(0, V)$
2	1.58125
4	1.5
8	1.375
16	1.36914..
32	1.34619..
256	1.3359375
1024	1.33398..

It is also not possible to explore fully the behavior of the ratio measures on $S_{k,m}$ as a function of k. For $m = 32$ and $b = 2$ Table 2 compares the value of $R^*(0, U_{k,32})$ to Rissanen's lower bound to that quantity, for those k values considered in [3].

TABLE 2		
k	$R^*(0, U_{k,32})$	Rissanen
2	1.50	1.45
4	1.33	1.27
5	1.29	1.23
6	1.23	1.20
12	1.16	1.11

Clearly $R^*(0, U_{m,m}) = h(q_m)/\log m \to 1$ as $m \to \infty$, and as noted in Section 3 the existence of asymptotically optimal codeword sets proves that $R^*(0, U_{k,\infty}) \to 1$ as $k \to \infty$.

Appendix: Specific algorithms for R^*, D^*, r^* and d^*.

A.1 Algorithms for R^* and D^*.

The function F required to compute D^* and R^* on $S_{k,m}$ is $j \log j$. The corresponding convex $\bigcup$ hull F' in (5.3) is

$$F'(j) = \begin{cases} j \log k, & 0 \le j \le k, \\ j \log j, & k < j < m+1, \end{cases} \tag{A.1}$$

and the difference function f in Problem 2 is given by

$$f(j) = \begin{cases} \log k, & 1 \le j \le k, \\ j \log j - (j-1)\log(j-1), & k < j < m+1. \end{cases} \tag{A.2}$$

If $k = m$ then $f(j) = \log k$ for all $j < m+1$. If $2 \le k < m$ then the inequalities

$$j \log(1 + 1/j) < \log e < (j+1)\log(1 + 1/j) \tag{A.3}$$

give

$$\log k \le f(j) \le \log je < f(j+1) < \log me, 1 \le j < m, \tag{A.4}$$

so that the number of evaluations of $t(k)$ required is at most

$$g(m) - g(1) = R(f(m) - f(1)) < R \log me/k, \tag{A.5}$$

and by (A.4) and the definition (5.8) of $t(k)$,

$$\begin{aligned} D + R \log je \le k \le D + R \log(j+1)e \\ \Rightarrow g(j) < k < g(j+2) \\ \Rightarrow j \le t(k) \le j+1, \end{aligned} \tag{A.6}$$

using D and R for x and y in Problem 2.

Using precomputed values of $g(1)$ and $g(m)$, the following algorithm finds $t(k)$ and makes at most one evaluation of $g(j)$ for some j so it requires $O(1)$ steps rather than the $O(\log m)$ steps required by a binary search.

$$\begin{aligned}
&\text{if } k \le g(1) \text{ then return } \mathbf{0};\\
&\text{if } k > g(m) \text{ then return } m;\\
&t := \lfloor b^{(k-D)/R}/e \rfloor;\\
&\text{if } k \le g(t+1) \text{ then return } t \text{ else return } t+1.
\end{aligned} \tag{A.7}$$

The first two lines follow from the definition of $t(k)$ in (5.8), and the rest from (A.6). It follows that n-bit values of $D^*(S_{k,m})$ and $R^*(S_{k,m})$ can be computed in $O(n \log me/k)$ steps. (Note that the algorithm terminates in its first or second line when $R = 0$ and $g(1) = g(m) = D$.)

A.2 Algorithms for r^* and d^* on $V_{k,m}$ for $k = b^a - (b-1)$.

The function $F(j)$ required to compute d^*, r^* and related cost measures on $V_{k,m}$ is $jh(q_j)$ in (4.11), and by Theorem 2 the relevant value of the parameter k in Problems 1 and 3 is b^a.

By Lemma 5 the corresponding lower convex hull $F'(j)$ is equal to $F(j)$ for j in J, where for finite m

$$\begin{aligned}
&J = \{0, b^a, b^{a+1}, \ldots, m-\rho, m\} \text{ and}\\
&\rho = \text{remainder of } m-1 \pmod{b-1},
\end{aligned} \tag{A.8}$$

and the other values of F' are given by linear interpolation between neighboring values in J; the difference function f in Problem 2 is

$$f(j) = \begin{cases} a, & 1 \le j \le b^a,\\ L(j-1) + b/(b-1), & b^a < j < m-\rho+1,\\ (F(m) - F(m-\rho))/\rho, & m-\rho < j < m+1. \end{cases} \tag{A.9}$$

Using d and r for x and y in Problem 2 the value of $t(k)$ in (5.8) is given by the following algorithm:

$$\begin{aligned}
&\text{if } k \le g(1) \text{ then return } \mathbf{0};\\
&\text{if } k > g(m) \text{ then return } m;\\
&\text{if } k > g(m-\rho) \text{ then return } m-\rho;\\
&\text{if } k \le g(b^a+1) \text{ then return } b^a;\\
&w := \lceil (k-d)/r - b/(b-1) \rceil;\\
&\text{return } b^w,
\end{aligned} \tag{A.10}$$

which runs in $O(1)$ steps, and at most $g(m) - g(1) \approx r \log m/k$ such evaluations are required to cover the integer slopes in the range of g. As in (A.7) the algorithm terminates in its first or second line when $r = 0$ and g(1)=g(m)=d.

A.3 Algorithms for d^* and r^* on $W_{k,m}$ for $k = b^a - (b-1)$.

The function $F(j)$ required to compute d^*, r^* and related cost measures on $W_{k,m}$ is $jh_m(q_j)$ in (4.11), and by Theorem 2 the relevant value of k for use in Problems 1 and 3 is $k = b^a - 1$.

For $b = 2$, by Lemma 5 the lower convex hull $F'(j)$ is equal to $F(j)$ for j in J, where

$$\begin{aligned} J &= \{0, 2^a - 1, 2^a, 2^{a+1}, \ldots, m_0, m\}, \\ m_0 &= 2^{L(m-1)} \text{ and} \\ L(j) &= \lfloor \log_2(j) \rfloor. \end{aligned} \tag{A.11}$$

The difference function f in Problem 2 is

$$f(j) = \begin{cases} a, & 1 \le j \le 2^a - 1, \\ a+1, & j = 2^a < m_0 + 1, \\ L(j) + 2, & 2^a < j < m_0 + 1, \\ (F(m) - F(m_0))/(m - m_0), & m_0 < j < m + 1. \end{cases} \tag{A.12}$$

Using d and r for x and y in Problem 2 the value of $t(k)$ in (5.8) is given by the following algorithm:

$$\begin{aligned} &\text{if } k \le g(1) \text{ then return } \mathbf{0}; \\ &\text{if } k > g(m) \text{ then return } m; \\ &\text{if } k > g(m_0) \text{ then return } m_0; \\ &\text{if } k \le g(2^a) \text{ then return } 2^a - 1; \\ &w := \lceil (k - d)/r - 2 \rceil; \\ &\text{return } 2^w. \end{aligned} \tag{A.13}$$

For $b > 2$ the lower convex hull F' of F is more complex. By Lemma 5 $F'(j) = F(j)$ for j in J, where

$$\begin{aligned} J &= \{0, b^a - 1, b^a + (b-2), b^{a+1} - 1, b^{a+1} + (b-2), \ldots, m_0, m\}, \\ m_0 &= \left\{ \begin{matrix} b^{L(m-1)} + b - 2 \text{ if } \rho \le 1, \\ m - \rho \text{ otherwise} \end{matrix} \right\}, \text{ and} \\ \rho &= \text{remainder of } m \pmod{b-1}, \end{aligned} \tag{A.14}$$

and its values at other arguments are given as before by interpolation. Then the function f in Problem 2 is

$$f(j) = \begin{cases} a, & 1 \le j \le b^a - 1, \\ L(j) + 1, & b^a \le j \le n(j) < n(m-1) + 1, \\ L(j) + b/(b-1), & b^a \le n(j) < j < m - \rho + 1, \\ (F(m) - F(m_0)))/(m - m_0)), & m_0 < j < m + 1. \end{cases} \tag{A.15}$$

The following algorithm finds $t(k)$ in $O(1)$ steps, using d and r for x and y in Problem 2.

if $k \leq g(1)$ then return 0;
if $k > g(m)$ then return m;
if $k > g(m_0)$ then return m_0;
if $k \leq g(b^a)$ then return $b^a - 1$; (A.16)
$v := (k-d)/r$; $w := \lceil v - 1 \rceil$;
if $v - w > 1/(b-1)$ then return $b^w - 1$;
else return $b^{w-1} + b - 2$.

REFERENCES

[1] Elias, P., "Minimum times and memories needed to compute the values of a function", J. Comput. Syst. Sci., vol.9 no.2, pp. 196-212, Oct. 1974.

[2] Elias, P., "Universal codeword sets and representations of the integers", IEEE Trans. on Information Theory vol. IT-21 no.2, pp. 194-203, March 1975.

[3] Rissanen, J., "Minimax codes for finite alphabets", IEEE Trans. on Information Theory vol. IT-24 no.3, pp.389-392, May 1978.

[4] Davisson, L.D. and Leon-Garcia, A., "A source matching approach to finding minimax codes", IEEE Trans. Information Theory vol. IT-26 no. 2, pp. 166-174, March 1980.

[5] Humblet, P.A., "Optimal source encoding for a class of integer alphabets", IEEE Trans. Information Theory vol. IT-24 no. 1, pp. 111-112, Jan. 1978.

[6] Levenstein, V.I., "The redundancy and deceleration of a separative encoding of the natural numbers", Probl. Cybern., no. 20, pp 173-179, Moscow, 1968.

[7] Davisson, L.D., "Universal noiseless coding", IEEE Trans. Information Theory vol. IT-19 no. 6, pp.783-795, Nov. 1973.

[8] Krichevsky, R.E. and Trofimov, V.K. "The performance of universal encoding", IEEE Trans. Information Theory vol. IT-27 no. 2, pp 199-207, March 1981.

[9] Shannon, C.E., "Prediction and entropy of printed English", Bell Syst. Tech J. vol. 30, pp. 50-64, Jan. 1951.