Pergamon

Appl. Math. Lett. Vol. 11, No. 5, pp. 121–126, 1998

0893-9659/98 $19.00 + 0.00

PII: S0893-9659(98)00091-3

Isoperimetric Theorems in the Binary Sequences of Finite Lengths

R. AHLSWEDE AND N. CAI
Fakulfät für Mathematik, Universität Bielefeld
Postfach 100131, D-33501 Bielefeld, Germany

(Received April 1997; accepted June 1997)

Abstract—We solve the isoperimetric problem for subsets in the set $\mathcal{X}^*$ of binary sequences of finite length for two cases:

(1) the distance counting the minimal number of insertions and deletions transforming one sequence into another;
(2) the distance, where in addition also exchanges of letters are allowed.

In the earlier work, the range of the competing subsets was limited to the sequences $\mathcal{X}^n$ of length n.

Keywords—Isoperimetry, Sequence spaces, Deletions, Insertions, Hamming distance, H^*-order.

1. THE PROBLEMS

The present note continues our paper [1]. We keep our earlier notation. Familiarity with [1] is not necessary but certainly helpful for an understanding of this paper.

We recall some definitions. For $\mathcal{X} = \{0,1\}$ and $n \in \mathbb{N}$, $\mathcal{X}^n$ denotes the space of binary sequences of length n. The fundamental object in our investigation is

$$\mathcal{X}^* = \bigcup_{n=0}^{\infty} \mathcal{X}^n,$$

the space of binary sequences of finite length. Here the sequence of length 0 is understood as the empty sequence ϕ.

Basic operations are deletions ∇ and insertions Δ. Here ∇ (respectively, Δ) means the deletion (respectively, insertion) of any letter.

We introduce again two distances, θ and δ, in $\mathcal{X}^*$. For x^m, $y^{m'} \in \mathcal{X}^*$, $\theta(x^m, y^{m'})$ counts the minimal number of insertions and deletions which transform one sequence into the other and $\delta(x^m, y^{m'})$ counts the minimal number of operations, if also exchanges of letters are allowed. For $\tau = \theta,\ \delta$, we define for $A \subset \mathcal{X}^*$

$$\Gamma_\tau^\ell(A) = \{x^{m'} :\ \text{there exists an } a^m \in A \text{ with } \tau\left(x^{m'}, a^m\right) \le \ell\}.$$

We abbreviate $\Gamma_\tau^1 = \Gamma_\tau$.

In [1], we showed that the initial segments of size u in Harper's order (introduced in [2]), or in short "the u^{th} initial segments in H-order" minimizes $|\Gamma_\theta^\ell(A)|$, $|\Gamma_\delta^\ell(A)|$, $|\Delta^\ell A|$, and $|\Delta^\ell A|$ for

Typeset by $\mathcal{A}_\mathcal{M}\mathcal{S}$-TEX

$A \subset \mathcal{X}^n$ with $|A| = u$, where $\Delta^\ell A$ is the subset of $\mathcal{X}^{n+\ell}$ obtained by inserting ℓ letters to the sequences in A and Δ^ℓ is defined analogously.

We introduce now $\Gamma^\ell_\Delta(A) = (\bigcup_{i=0}^{\ell} \Delta^i A)$ $(\Gamma^1_\Delta = \Gamma_\Delta)$.

In this note, we change the range of A from subsets of $\mathcal{X}^n$ to subsets of $\mathcal{X}^*$.

Clearly,

$$\Gamma^\ell_\Delta(A) \subset \Gamma^\ell_\theta(A) \subset \Gamma^\ell_\delta(A), \qquad \text{for all } A \subset \mathcal{X}^*. \tag{1.1}$$

The role of the H-order for the former problems in [1] for the new isoperimetric problems is played by what we call H^*-order. Its definition follows next.

2. THE H^*-ORDER

Recalling that x^n precedes y^n in the squashed order on $\{x^n \in \mathcal{X}^n : \sum_{i=1}^n x_i = k\}$ exactly if $x_t < y_t$, if t is the largest number s with $x_s \neq y_s$, and that x^n precedes y^n in the H-order on $\mathcal{X}^n$, exactly if $\sum_{t=1}^n x_t < \sum_{t=1}^n y_t$ or $\sum_{t=1}^n x_t = \sum_{t=1}^n y_t$ and $(1-x_1, \dots, 1-x_n)$ precedes $(1-y_1, \dots, 1-y_n)$ in the squashed order, we introduce the following H^*-order. For x^n, $x^{m'} \in \mathcal{X}^*$, x^m precedes $y^{m'}$, exactly if $m < m'$ or $m = m'$ and x^m precedes $y^{m'}$ in the H-order.

Katona [3] has shown that for any integers n and $u \in [0, 2^n]$ there is a unique binomial representation

$$u = \binom{n}{n} + \cdots + \binom{n}{k+1} + \binom{\alpha_k}{k} + \cdots + \binom{\alpha_t}{t} \tag{2.1}$$

(with $n > \alpha_k > \cdots > \alpha_t \geq t \geq 1$). He introduced the function

$$G(n,u) = \binom{n}{n} + \cdots + \binom{n}{k+1} + \binom{n}{k} + \binom{\alpha_k}{k-1} + \cdots + \binom{\alpha_t}{t-1}, \tag{2.2}$$

and proved that for $0 \leq u_1 \leq u_0$ and $u \leq u_0 + u_1$,

$$G(n,u) \leq \max(u_0, G(n-1,u_1)) + G(n-1,u_0). \tag{2.3}$$

It immediately follows from the uniqueness of the representation (2.1) that every positive integer N can be uniquely represented as

$$\begin{aligned} N &= 1 + 2 + \cdots + 2^{n-1} + \binom{n}{n} + \cdots + \binom{n}{k+1} + \binom{\alpha_k}{k} + \cdots + \binom{\alpha_t}{t} \\ &= 1 + 2 + \cdots + 2^{n-1} + u \, (0 \leq u < 2^n \text{ and } u \text{ as in } (2.1)). \end{aligned} \tag{2.4}$$

We introduced in [1] (for u as in (2.1))

$$\overset{\Delta}{G}(n,u) = \binom{n+1}{n+1} + \cdots + \binom{n+1}{k+1} + \binom{\alpha_k+1}{k} + \cdots + \binom{\alpha_t+1}{t}, \tag{2.5}$$

and proved (in Lemma 6) that ΔS is the $\overset{\Delta}{G}(n,|S|)^{\text{th}}$ initial segment in the H-order on $\mathcal{X}^{n+1}$, if S is an initial segment in the H-order on $\mathcal{X}^n$.

Consequently, by the definition of our H^*-order,

$$\Gamma_\Delta(S') = \Gamma_\theta(S') = \Gamma_\delta(S') \tag{2.6}$$

is the $G^*(N)^{\text{th}}$ initial segment in the H^*-order on $\mathcal{X}^*$, for the N^{th} initial segment S' in the H^*-order on $\mathcal{X}^*$, if we introduce

$$G^*(N) = 1 + 2 + \cdots + 2^{n-1} + 2^n + \overset{\Delta}{G}(n,u) = (2^{n+1} - 1) + \overset{\Delta}{G}(n,u) \text{ (for } N \text{ in (2.4))}. \tag{2.7}$$

By (2.1), (2.4), and (2.5) (see, also, [1]),

$$G(n,u) + u = \overset{\Delta}{G}(n,u), \tag{2.8}$$

and therefore, (2.7) imply that

$$G^*(N) = N + 2^n + G(n,u). \tag{2.9}$$

3. THE RESULTS

THEOREM 1. *For all $A \subset \mathcal{X}^*$ with $|A| = N$,*

$$G^*(N) \le |\Gamma_\Delta(A)| \le |\Gamma_\theta(A)| \le |\Gamma_\delta(A)|, \tag{3.1}$$

and all inequalities in (3.1) are equalities, if A is an initial segment in H^*-order on $\mathcal{X}^*$.

If S is an initial segment in the H^*-order, then so is $\Gamma_\Delta(S) = \Gamma_\theta(S) = \Gamma_\delta(S)$.

Therefore, Theorem 1 can be applied repeatedly and gives our general isoperimetric inequalities.

THEOREM 2. *For every integer $N \in \mathbb{N}$, S_N, the N^{th} initial segment in H^*-order has for every $\ell \in \mathbb{N}$, the same ℓ^{th} boundaries in all three cases, that is,*

$$\Gamma_\Delta^\ell(S_N) = \Gamma_\theta^\ell(S_N) = \Gamma_\delta^\ell(S_N),$$

and they are minimal among sets of cardinality N, that is,

$$\left|\Gamma_\Delta^\ell(S_N)\right| = \min_{A\subset\mathcal{X}^*,\ |A|=N} \left|\Gamma_\Delta^\ell(A)\right| = \min_{A\subset\mathcal{X}^*,\ |A|=N} \left|\Gamma_\tau^\ell(A)\right|, \qquad \tau = \theta, \delta. \tag{3.2}$$

4. TWO AUXILIARY RESULTS

To prove Theorem 1, we need the following inequalities.

LEMMA 1. *For $0 \le N_1 \le N_0$,*

$$G^*(N_0 + N_1 + 1) \le \max(N_0 + N_1 + 1, G^*(N_1)) + G^*(N_0) + 1.$$

LEMMA 2. *For $0 \le N_1 \le N_0$,*

$$G^*(N_0 + N_1) \le \max(N_0 + N_1, G^*(N_1)) + G^*(N_0).$$

In the proofs in Sections 5 and 6, we use simple properties of the function G.

PROPOSITION. *For $u \in [0, 2^n]$ and $n \in \mathbb{N}$, G is nondecreasing in u and*

$$G(n,u) \le 2^n. \tag{4.1}$$

Here, equality holds exactly if

$$u > 2^n - n - 1, \tag{4.2}$$

$$u < G(n,u), \qquad (\text{for } 2^n > u > 0)\,, \tag{4.3}$$

and

$$G(n,u) \le u + G(n-1,u). \tag{4.4}$$

PROOF. Here (4.4) follows from (2.3), for $u_1 = 0$ and $u = u_0$. The other statements follow readily with definition (2.2).

The reader, who believes Lemmas 1 and 2, can immediately continue with Section 7.

5. PROOF OF LEMMA 1

Let $0 \le N_1 \le N_0$ and

$$N = N_0 + N_1 + 1 = 1 + \cdots + 2^{n-1} + u = 2^n - 1 + u, \qquad (0 \le u < 2^n), \tag{5.1}$$

then $2^{n-1} - 1 \le N_0 < 2^{n+1} - 1$.

CASE 1.

$$2^{n-1} - 1 \le N_1 \le N_0 < 2^n - 1. \tag{5.2}$$

Here we can write

$$N_0 = 1 + 2 + \cdots + 2^{n-2} + u_0 = 2^{n-1} - 1 + u_0, \qquad (0 \le u_0 < 2^{n-1}), \tag{5.3}$$

and

$$N_1 = 1 + 2 + \cdots + 2^{n-2} + u_1 = 2^{n-1} - 1 + u_1, \qquad (0 \le u_1 \le u_0). \tag{5.4}$$

By (5.1), (5.3), and (5.4), we have that

$$u = u_0 + u_1. \tag{5.5}$$

Thus, it follows from (5.3), (5.4), (2.9), (2.3), and (5.1) that the RHS in Lemma 1 equals $\max(u_0, G(n-1, u_1)) + (N_1 + 2^{n-1}) + N_0 + 2^{n-1} + G(n-1, u_0) + 1$ (by (5.3), (5.4), and (2.9)) $\ge G(n, u_0 + u_1) + (N_0 + N_1 + 1) + 2^n$ (by (2.3)) $=$ LHS in Lemma 1 (by (2.9) and (5.1)).

CASE 2.

$$N_0 \ge 2^n - 1. \tag{5.6}$$

Here we write

$$N_0 = 1 + \cdots + 2^{n-1} + u_0, \qquad (0 \le u_0 < 2^n). \tag{5.7}$$

Thus by (5.1), (5.7), (2.9), (5.6), (4.1), and (5.1), RHS in Lemma 1 $\ge N + N_0 + 2^n + G(n, u_0) + 1$ (by (5.1), (5.7), and (2.9)) $\ge N + 2^{n+1} + G(n, u_0)$ (by (5.6)) $\ge N + 2^{n+1} \ge N + 2^n + G(n, u)$ (by (4.1)) $=$ LHS in Lemma 1 (by (5.1) and (2.9)).

CASE 3.

$$N_1 < 2^{n-1} - 1 \le N_0 < 2^n - 1. \tag{5.8}$$

Here (5.3) holds, and by (5.1), (5.3), and (5.8),

$$u_0 = N - N_1 - 1 - (2^{n-1} - 1) > N - 2 \cdot (2^{n-1} - 1) - 1 = u + (2^n - 1) - 1 - 2^n + 2 = u. \tag{5.9}$$

So, we have, by (5.1), (5.3), (2.9), (5.9), and (4.4) that RHS in Lemma 1 $\ge N + N_0 + 2^{n-1} + G(n-1, u_0) + 1$ (by (5.1), (5.3), and (2.9)) $= N + 2^n + u_0 + G(n-1, u_0)$ (by (5.3)) $> N + 2^n + u + G(n-1, u)$ (by (5.9)) $\ge N + 2^n + G(n, u)$ (by (4.4)) $=$ LHS in Lemma 1 (by (5.1) and (2.9)).

6. PROOF OF LEMMA 2

Let $0 \le N_1 \le N_0$ and

$$N' = N_0 + N_1 = 1 + 2 + \cdots + 2^{n-1} + u' = 2^n - 1 + u' \, (0 \le u' < 2^n), \tag{6.1}$$

then $2^{n-1} \le N_0 < 2^{n+1} - 1$.

CASE 1. EQUATION (5.2) HOLDS. Then, also (5.3),(5.4) hold, and

$$u' + 1 = u_0 + u_1. \tag{6.2}$$

Similarly, as in the same case in the proof of Lemma 1, we have now by (5.3), (5.4), (2.9), and (6.1), that the RHS in Lemma 2

$$\begin{aligned}&= \max(u_0 - 1, G(n-1, u_1)) + N_1 + 2^{n-1} + N_0 + 2^{n-1} + G(n-1, u_0) \\ &= \max(u_0 - 1, G(n-1, u_1)) + N' + 2^n + G(n-1, u_0),\end{aligned} \tag{6.3}$$

which together with (6.2), (2.3), (2.9), and (6.1) implies Lemma 2 for $u_1 \le u_0 - 1$, since $G(n-1, u_0 - 1) \le G(n-1, u_0)$.

Otherwise, $u_1 = u_0$, and therefore, by (4.3)

$$u_0 - 1 < u_0 \le G(n-1, u_1). \tag{6.4}$$

Thus, by (6.2), (2.3), and (6.1), again RHS of (6.3) $= \max(u_0, G(n-1, u_1)) + N' + 2^n + G(n-1, u_0) \ge N' + 2^n + G(n, u_0 + u_1) \ge N' + 2^n + G(n, u') =$ LHS in Lemma 2.

CASE 2. EQUATION (5.6) HOLDS. Hence, also (5.7) holds. By (6.1), (5.7), and (2.9),

$$\text{RHS of Lemma 2} \ge N' + N_0 + 2^n + G(n, u_0) \ge N' + 2^{n+1} - 1 + u_0 + G(n, u_0). \tag{6.5}$$

By (6.1), (4.1), and (2.9), the RHS in (6.5) is not smaller than the LHS in Lemma 2 unless $u_0 = 0$ and $G(n, u') = 2^n$.

Assume that $u_0 = 0$ and $G(n, u') = 2^n$. Then by (4.1) and (4.2), $u' > 2^n - n - 1$. So, in this case, by (5.7) and (6.1),

$$N_1 = N' - N_0 = u' - u_0 > 2^n - n - 1. \tag{6.6}$$

This implies that N_1 can be represented as

$$N_1 = 1 + 2 + \cdots + 2^{n-2} + u_1, u_1 > 2^{n-1} - n \left(= 2^{n-1} - (n-1) - 1\right). \tag{6.7}$$

Then, by (6.7), (5.7), (6.1), (2.9), (4.1), and (4.2), we have RHS Lemma 2 $\ge N_1 + 2^{n-1} + G(n-1, u_1) + N_0 + 2^n + G(n, u_0) = N' + 2^{n+1} =$ LHS in Lemma 2, again.

CASE 3. EQUATION (5.8) HOLDS. Here, similarly to (5.9), by (6.1) and (5.8), we have that

$$u_0 = N' - N_1 - \left(2^{n-1} - 1\right) = (2^n - 1) + u' - N_1 - \left(2^{n-1} - 1\right) > u' + 1. \tag{6.8}$$

Thus, since $G(n-1, \cdot)$ is nonincreasing, by (2.9), (6.8), and (4.4), RHS in Lemma 2 $\ge N' + N_0 + 2^{n-1} + G(n-1, u_0) = N' + (2^{n-1} - 1) + u_0 + 2^{n-1} + G(n-1, u_0) \ge N' + 2^n + (u_0 - 1) + G(n-1, u_0 - 1) \ge$ LHS in Lemma 2.

7. PROOF OF THEOREM 1

By (1.1) and (2.6), it is sufficient to show that for all $A \subset \mathcal{X}^*$ with $|A| = N$,

$$G^*(N) \le |\Gamma_\Delta(A)|. \tag{7.1}$$

We show it by induction on N. For $N = 1$, (7.1) obviously holds.

For $B \subset \mathcal{X}^*$ and $i = 0, 1$, we define

$$B_i = \{(b_1, \ldots, b_\ell) : (b_1, \ldots, b_\ell, i) \in B\}, \tag{7.2}$$

$$B * i = \{(b_1, \ldots, b_m, i) : (b_1, \ldots, b_m) \in B\}, \tag{7.3}$$

and

$$\hat{B}_i = \{(b_1, \ldots, b_j) : B_j = i \text{ and } (b_1, \ldots, b_j) \in B\}. \tag{7.4}$$

Now fix $A \subset \mathcal{X}^*$ and assume w.l.o.g. that $|\hat{A}_1| \le |\hat{A}_0|$. Write $|\hat{A}_i| = N_i$ for $i = 0, 1$. With these notions, if $N_0 \ne N$, then

$$\left|\left(\widehat{\Gamma_\Delta A}\right)_i\right| \ge \max\left(N, G^*(N_i)\right), \qquad \text{for } i = 0, 1, \tag{7.5}$$

because $A * i \subset (\widehat{\Gamma_\Delta A})_i$, $(\Gamma_\Delta A_i) * i \subset (\widehat{\Gamma_\Delta A})_i$ and by the induction hypothesis $|\Gamma_\Delta A_i| \ge G^*(N_i)$.

CASE 1. $\phi \in A$. Then,

$$N = |A| = N_0 + N_1 + 1 \tag{7.6}$$

and

$$\Gamma_\Delta(A) = \left(\widehat{\Gamma_\Delta(A)}\right)_0 \cup \left(\widehat{\Gamma_\Delta A}\right)_1 \cup \{\phi\}. \tag{7.7}$$

Thus by (7.5),

$$|\Gamma_\Delta(A)| \ge \max\left(N_0 + N_1 + 1, G^*(N_1)\right) + G^*(N_0) + 1.$$

Therefore, Theorem 1 follows from Lemma 1 in this case.

CASE 2. $\phi \notin A$. Then

$$N = N_0 + N_1, \tag{7.8}$$

and we can assume that $N_0 \ne N_1$, because otherwise we can replace A by A_0 without changing the size of the set, and this change does not increase the size of "Γ_Δ". We are now able to use (7.5) to obtain that

$$|\Gamma_\Delta(A)| \ge \max\left(N_0 + N_1, G^*(N_1)\right) + G^*(N_0),$$

because in this case $\Gamma_\Delta(A) = (\widehat{\Gamma_\Delta A})_0 \cup (\widehat{\Gamma_\Delta A})_1$. Finally, Theorem 1 follows from Lemma 2.

REMARK. Inspection of the proof of the theorem shows that initial segments in H^*-order may not be the only minimal sets (of course in the isomorphic sense) for which we have equality in Lemma 2 in our "extremal problems of Γ_Δ". Indeed, when $|A| = N = 4$, $G^*(4) = 11$, the 4^{th} initial segment in the H^*-order is $S = \{\phi, 0, 1, 00\}$ and $\Gamma_\Delta(S)$ contains 11 sequences, namely, $\phi, 0, 1, 00, 01, 10, 11, 000, 001, 010$, and 100. If $N_0 = 3$ and $N_1 = 1$, then both sides in Lemma 2 equal 11. If $A = \{0, 00, 01, 10\}$, then $\Gamma_\Delta(A)$ contains $0, 00, 01, 10, 000, 001, 010, 100, 011, 101$, and 110, that is also 11 sequences. This example shows that Lemma 2 is really necessary.

REFERENCES

1. R. Ahlswede and N. Cai, Shadows and isoperimetry under the sequence-subsequence relation, *Combinatorica* **17** (1), 11–29, (1997).
2. L.H. Harper, Optimal numberings and isoperimetric problems on graphs, *J. Combin. Theory* **1**, 385–393, (1966).
3. G. Katona, The Hamming shere has minimum boundary, *Studia Sci. Math. Hungar.* **10**, 131–140, (1975).