Discrete Mathematics 49 (1984) 1–5
North-Holland

A TWO-FAMILY EXTREMAL PROBLEM IN HAMMING SPACE

Rudolf AHLSWEDE*

Faculty of Mathematics, University of Bielefeld, Universitätsstrasse 1, 4800 Bielefeld, W. Germany

Abbas EL GAMAL†, King F. PANG‡

Information Systems Laboratory, Electrical Engineering Department, Stanford University, Stanford, CA 94305, USA

Received 15 March 1983

The pair (A, B); $A, B \subseteq \{0,1\}^m$; is called an (m, δ)-system, if for the Hamming distance function d,

$$d(a, b) = \delta \ \forall a \in A, \forall b \in B. \tag{1}$$

Let S_δ^m denote the set of those systems. We consider the function

$$M(m, \delta) \triangleq \max\{|A|\,|B| : (A, B) \in S_\delta^m\} \tag{2}$$

and prove the following Theorem: For $m = 1, 2, \ldots$, $\max_{0 \le \delta \le m} M(m, \delta) = 2^{2n}$, if $m = 2n$ or $m = 2n+1$. The maximum is assumed for $\delta = n$.

1. Introduction

The study of (m, δ)-systems is motivated by the problem of lower bounding the two-way complexity (in the sense of Yao [1]) of the Hamming distance function. Results in this direction will be contained in [2].

Two-family extremal problems have frequently been considered in the literature [3]. Replacement of (1) by

$$d(a, b) \geqslant \delta \quad \forall a \in A, \quad \forall b \in B \tag{1'}$$

yields an extremal problem, which has been solved in [4]. However, in spite of the similarity between conditions (1) and (1'), the present proof techniques are quite different from those in [4]. Actually, we give two proofs of the Theorem. The first is by a 1-step and the second by a 2-step induction in m. The examples

$$A \triangleq \{01, 10\}^n, \qquad B \triangleq \{11, 00\}^n \tag{3}$$

$$A \triangleq \{01, 10\}^n \times \{0\}, \qquad B \triangleq \{11, 00\}^n \times \{0\} \tag{4}$$

are crucial for understanding the Theorem. They immediately yield

* Work done while Visiting Professor at ISL, Stanford.

† Work partially supported by DARPA under Contract MDA-0680 and by U.S. Air Force under Contract F49620-79C-0058.

‡ Work partially supported by NSF under Contract 80-26102.

Lemma 1.

$$M(2n, n) \geq 2^{2n}, \quad M(2n+1, n) \geq 2^{2n} \qquad (n = 1, 2, \ldots)$$

Thus only the inequality

$$\max_{0 \leq \delta \leq m} M(m, \delta) \leq 2^{2n} \tag{I}$$

remains to be proved.

First we show that this inequality follows from either one of the following two propositions. Actually, these derivations establish also their equivalence. The proofs for the propositions will be given subsequently.

Proposition 1. $M(2n+1, n) = M(2n, n)$, $(n = 1, 2, \ldots)$.

Proposition 2. $M(2n, n) \leq 2^{2n}$, $(n = 1, 2, \ldots)$.

2. Preliminaries

The operation $\bar{}$ applied to a sequence denotes complementation, that is, component-wise exchange of 0's and 1's. When applied to a set of sequences, it is understood in the *Minkowski sense*. For ease of reference, a simple property of the Hamming distance function with respect to complementation is stated as

Lemma 2. (i) $d(\bar{a}, b) + d(a, b) = m$, $d(\bar{a}, \bar{b}) = d(a, b)$, $(a, b \in \{0, 1\}^m)$,
(ii) $(A, B) \in S_\delta^m \Rightarrow (A, B) \in S_{m-\delta}^m$,
(iii) $M(m, \delta) = M(m, m - \delta)$.

We also adopt the following notation: For a set $C \subset \{0, 1\}^m$ and $\varepsilon \in \{0, 1\}$ define

$$C_\varepsilon^t \triangleq \{(c_1, \ldots, c_m) \in C : c_t = \varepsilon\} \subset \{0, 1\}^m \tag{5}$$

$$C_\varepsilon^{*t} \triangleq \{(c_1, \ldots, c_{t-1}, c_{t+1}, \ldots, c_m) : (c_1, \ldots, c_{t-1}, \varepsilon, c_{t+1}, \ldots, c_m) \in C\} \subset \{0, 1\}^{m-1} \tag{6}$$

Analogously, for two components s, t we define $C_{\varepsilon\eta}^{st} \subset \{0, 1\}^m$ and $C_{\varepsilon\eta}^{*st} \subset \{0, 1\}^{m-2}$.

3. Proposition 1 $\Rightarrow$ Proposition 2

We proceed by induction in n. The case $n = 1$ is settled by inspection. Now for $(A, B) \in S_{n+1}^{2(n+1)}$, clearly

$$A = A_1^1 \cup A_0^1, \qquad B = B_1^1 \cup B_0^1 \tag{7}$$

and for $\varepsilon \in \{0, 1\}$

$$|A_\varepsilon^{*1}| = |A_\varepsilon^1|, \qquad |B_\varepsilon^{*1}| = |B_\varepsilon^1|; \tag{8}$$

$$|A_\varepsilon^{*1}|\,|B_\varepsilon^{*1}| \leq M(2n+1, n+1); \tag{9}$$

$$|A_\varepsilon^{*1}|\,|B_\varepsilon^{*1}| \leq M(2n+1, n). \tag{10}$$

Since by Lemma 2, $M(2n+1, n+1) = M(2n+1, n)$, the relations (7)–(10) and Proposition 1 imply

$$|A|\,|B| = \sum_{\varepsilon,\eta \in \{0,1\}} |A_\varepsilon^1|\,|B_\eta^1| \leq 4M(2n, n). \tag{11}$$

Since $M(2n, n) \leq 2^{2n}$ by hypothesis, thus $M(2(n+1), n+1) \leq 4(2^{2n}) = 2^{2(n+1)}$. □

4. Proposition 2 ⇒ (I)

Case $m = 2n$. For $(A, B) \in S_\delta^{2n}$, consider $(A \times \bar{A}, B \times \bar{B})$. By Lemma 2, this is an element of S_{2n}^{4n}. Therefore, by Proposition 2, $|A \times \bar{A}|\,|B \times B| = (|A|\,|B|)^2 \leq 2^{4n}$ and hence $M(2n, \delta) \leq M(2n, n)$.

Case $m = 2n+1$. Since $d(\bar{a}, b) + d(a, b) = m$ and m is odd, necessarily $d(\bar{a}, b) \neq d(a, b)$ and thus for $(A, B) \in S_\delta^m$ also $A \cap \bar{A} = \phi$, $B \cap \bar{B} = \emptyset$. By Lemma 2,

$$(C, D) \triangleq (A \times \bar{A} \cup \bar{A} \times A,\ B \times B \cup \bar{B} \times \bar{B}) \in S_{2n+1}^{2(2n+1)}$$

and thus by Proposition 2 $|C|\,|D| = 4(|A|\,|B|)^2 \leq 2^{2(2n+1)}$, which gives $|A|\,|B| \leq 2^{2n}$. Hence, we have derived (I) and in conjunction with Lemma 1, Proposition 1 is also proved.

5. Proof of Proposition 1

The proof is based on two key observations. For any $(A, B) \in S_n^{2n+1}$:

$$\forall t \in \{1, \ldots, 2n+1\} \quad \text{if} \quad A_\varepsilon^{*t} \cap \overline{A_\varepsilon^{*t}} \neq \emptyset \quad \text{then } B_{\bar{\varepsilon}}^{*t} = \emptyset. \tag{OI}$$

Clearly for $a, \bar{a} \in A_\varepsilon^{*t}$ and $b \in B_{\bar{\varepsilon}}^{*t}$ we have $d(a, b) = d(\bar{a}, b) = n-1$ in contradiction to $d(\bar{a}, b) = 2n - d(a, b) = n+1$.

$$\exists t \in \{1, \ldots, 2n+1\}: |A_1^t|\,|B_1^t| + |A_0^t|\,|B_0^t| \geq |A_1^t|\,|B_0^t| + |A_0^t|\,|B_1^t|. \tag{OII}$$

For this just notice that $\sum_{t=1}^{2n+1} (|A_1^t|\,|B_1^t| + |A_0^t|\,|B_0^t|)$ counts the number of identical components for *all* pairs of sequences $(a, b) \in A \times B$ and therefore equals $(n+1)\,|A|\,|B|$. On the other hand $\sum_{t=1}^{2n+1} (|A_1^t|\,|B_0^t| + |A_0^t|\,|B_1^t|)$ counts the number of *distinct* components for all pairs of sequences $(a, b) \in A \times B$ and therefore equals $n\,|A|\,|B|$, a smaller quantity. The Pigeon Hole Principle gives (OII).

We can assume without loss of generality that $t=2n+1$ and *omit the index t*. Notice that (A_1, B_1) and (A_0, B_0) are $(2n, n)$-systems.

Case 1. $\exists \varepsilon \in \{0,1\}$ with $A_\varepsilon^* \cap \overline{A_\varepsilon^*} \neq \emptyset$ *and* $B_\varepsilon^* \cap \overline{B_\varepsilon^*} \neq \emptyset$. By (OI), $A_{\bar\varepsilon}^* = B_{\bar\varepsilon}^* = \emptyset$ and thus $|A|\,|B| = |A_\varepsilon^*|\,|B_\varepsilon^*| \leq M(2n, n)$.

Case 2. $\exists \varepsilon \in \{0,1\}$ with $A_\varepsilon^* \cap \overline{A_\varepsilon^*} \neq \emptyset$ *and* $B_\varepsilon^* \cap \overline{B_\varepsilon^*} = \emptyset$ (resp. vice versa). By (OI) $B_{\bar\varepsilon}^* = \emptyset$ and thus $|A|\,|B| \leq |A_\varepsilon^*|\,|B_\varepsilon^*| + |A_{\bar\varepsilon}^*|\,|B_\varepsilon^*| \leq 2\,|A_\varepsilon^*|\,|B_\varepsilon^*|$ (by OII)). Replace now B_ε^* by $D \triangleq B_\varepsilon^* \cup \overline{B_\varepsilon^*}$. Since $(A_\varepsilon^*, D) \in S_n^{2n}$, we get again $|A|\,|B| \leq |A_\varepsilon^*|\,|D| \leq M(2n, n)$.

Case 3. $\forall \varepsilon \in \{0,1\}$: $A_\varepsilon^* \cap \overline{A_\varepsilon^*} = \emptyset$, $B_\varepsilon^* \cap \overline{B_\varepsilon^*} = \emptyset$. Choose now ε such that $|A_\varepsilon^*|\,|B_\varepsilon^*| \geq |A_{\bar\varepsilon}^*|\,|B_{\bar\varepsilon}^*|$ and define $C \triangleq A_\varepsilon^* \cup \overline{A_\varepsilon^*}$, $D \triangleq B_\varepsilon^* \cup \overline{B_\varepsilon^*}$. Now $(C, D) \in S_n^{2n}$ and

$$\begin{aligned} |A|\,|B| &= |A_1^*|\,|B_1^*| + |A_0^*|\,|B_0^*| + |A_1^*|\,|B_0^*| + |A_0^*|\,|B_1^*| \\ &\leq 2(|A_1^*|\,|B_1^*| + |A_0^*|\,|B_0^*|) \quad \text{(by OII)} \\ &\leq 4\,|A_\varepsilon^*|\,|B_\varepsilon^*| \quad \text{(by choice of } \varepsilon) \\ &= |C|\,|D| \leq M(2n, n). \end{aligned}$$

6. Proof of Proposition 2

Again the proof is based on two observations.

If $(A, B) \in S_n^{2n}$, then $(\bar A, B), (A, \bar B), (\bar A, \bar B) \in S_n^{2n}$ and also $(A \cup \bar A, B \cup \bar B) \in S_n^{2n}$. We can therefore assume $A = \bar A$, $B = \bar B$ and thus

$$|A_\varepsilon^t| = |A_{\bar\varepsilon}^t| = \tfrac{1}{2}|A| \quad \text{and} \quad |B_\varepsilon^t| = |B_{\bar\varepsilon}^t| = \tfrac{1}{2}|B| \qquad (1 \leq t \leq 2n,\ \varepsilon \in \{0,1\}). \tag{OIII}$$

Further

$$\exists t \in \{2, 3, \ldots, 2n\} \quad \text{and} \quad \exists \eta \in \{0,1\}: |A_{1\eta}^{1t}|\,|A_1^1|^{-1} \geq \tfrac{1}{2} \quad \text{and} \quad |B_{1\bar\eta}^{1t}|\,|B_1^1|^{-1} \geq \tfrac{1}{2}. \tag{OIV}$$

For this, notice that $\sum_{t=2}^{2n} |A_{10}^{1t}|\,|B_{11}^{1t}| + |A_{11}^{1t}|\,|B_{10}^{1t}|$ counts the number of *distinct* components for *all* pairs of sequences $(a, b) \in A_1^1 \times B_1^1$ and therefore equals $n\,|A_1^1|\,|B_1^1|$. Again by the Pigeon Hole Principle there exists a $t \in \{2, 3, \ldots, 2n\}$ with $|A_{10}^{1t}|\,|B_{11}^{1t}| + |A_{11}^{1t}|\,|B_{10}^{1t}| \geq (n/(2n-1))\,|A_1^1|\,|B_1^1| > \frac{1}{2}|A_1^1|\,|B_1^1|$. This implies (OIV).

Now again we distinguish among three cases.

Case 1. $A_{1\eta}^{*1t} \cap \overline{A_{1\eta}^{*1t}} \neq \emptyset$ *and* $B_{1\bar\eta}^{*1t} \cap \overline{B_{1\bar\eta}^{*1t}} \neq \emptyset$. Notice that by the distance properties necessarily, $A_{1\bar\eta}^{*1t} = \emptyset$ and $B_{1\eta}^{*1t} = \emptyset$. By (OIII) therefore $|A|\,|B| = 4\,|A_{1\eta}^{*1t}|\,|B_{1\bar\eta}^{*1t}| \leq 4M(2(n-1), n-1)$.

Case 2. $A_{1\eta}^{*1t} \cap \overline{A_{1\eta}^{*1t}} \neq \emptyset$ *and* $B_{1\bar\eta}^{*1t} \cap \overline{B_{1\bar\eta}^{*1t}} = \emptyset$. By the previous argument necessarily, $B_{1\eta}^{*1t} = \emptyset$. Define now $C \triangleq A_{1\eta}^{*1t}$ and $D \triangleq B_{1\bar\eta}^{*1t} \cup \overline{B_{1\bar\eta}^{*1t}}$. Since $d(a, b) = d(\bar a, b)$ and obviously $d(\bar a, \bar b) = d(a, b)$, also $d(a, \bar b) = d(\bar a, b)$. Thus $(C, D) \in S_{n-1}^{2n-2}$ and, by (OIII) and (OIV), $|A|\,|B| \leq 4\,|C|\,|D| \leq 4M(2n-2, n-1)$.

Case 3. $A_{1\eta}^{*1t} \cap \overline{A_{1\eta}^{*1t}} = \emptyset$ *and* $B_{1\bar\eta}^{*1t} \cap \overline{B_{1\bar\eta}^{*1t}} = \emptyset$. Since $\bar{\bar A} = A$ and $\bar{\bar B} = B$ we have now $A_{0\bar\eta}^{*1t} = \overline{A_{1\eta}^{*1t}}$, $B_{0\eta}^{*1t} = \overline{B_{1\bar\eta}^{*1t}}$. Furthermore, since $d(1\eta, 1\bar\eta) = d(0\bar\eta, 0\eta) = 1$, for

$C \triangleq A^{*1t}_{1\eta} \cup \overline{A^{*1t}_{1\eta}}$, $D \triangleq B^{*1t}_{1\bar{\eta}} \cup \overline{B^{*1t}_{1\bar{\eta}}}$ we have $(C, D) \in S^{2n-2}_{n-1}$ and thus again, by (OIII) and (OIV), $|A||B| \leq 4|C||D| \leq 4M(2n-2, n-1)$.

By the inductive hypothesis for $n-1$, $M(2n-2, n-1) \leq 2^{2(n-1)}$ and hence $M(2n, n) \leq 4(2^{2(n-1)}) = 2^{2n}$.

References

[1] A.C. Yao, Some complexity questions related to distributive computing, 11th ACM Symp. on Theory of Computing, (1979) 209–213.
[2] A. El Gamal and K.F. Pang, On the communication complexity of computation, in preparation.
[3] C. Greene and D.J. Kleitman, Proof techniques in the theory of finite sets, MIT Lecture Notes.
[4] R. Ahlswede and G. Katona, Contributions to the geometry of Hamming spaces, Discrete Math. 17 (1977) 1–22.