Discrete Mathematics 58 (1986) 191–193
North-Holland

NOTE

EXPLICIT CONSTRUCTION OF EXPONENTIAL SIZED FAMILIES OF k-INDEPENDENT SETS

N. ALON*

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Received 22 October 1984

Error correcting codes are used to describe explicit collections F_k of subsets of $\{1, 2, \ldots n\}$, with $|F_k| > 2^{c_k \cdot n}$ $(c_k > 0)$, such that for any selections A, B of k_1 and k_2 of members of F_k with $k_1 + k_2 = k$, there are elements in all the members of A and not in the members of B. This settles a problem of Kleitman and Spencer and a similar problem of Kleitman, Shearer and Sturtevant.

1. Introduction

A collection F of subsets of $N = \{1, 2, \ldots, n\}$ is *k-independent* if for every k distinct members $A_1, A_2, \ldots, A_k$ of F all 2^k intersections $\bigcap_{j=1}^{k} B_j$ are nonempty, where each B_j can be either A_j or its complement $\bar{A}_j$. Kleitman and Spencer [4] proved that for every fixed k there exists a k-independent collection F_k on n elements of size $\geq 2^{c_k \cdot n}$, where $c_k > 0$ is independent of n. Their proof is nonconstructive, i.e., it gives no explicit construction of such an F_k. They thus raised the problem of finding an explicit construction of such a collection, and mention that they have not been able to find one even for $k = 3$.

Here we settle their problem by applying the theory of error correcting codes to derive explicit constructions of k-independent families on n elements of size exponential in n. Specifically, we use a recent construction of Friedman [3] of certain generalized Justesen codes to describe the required families.

Our method also settles another problem, raised by Kleitman, Shearer and Sturtevant in [5]. They proved that the maximal cardinality of a collection of n-element sets with the property that the intersection of no two is included in a third is exponential in n, and raised the problem of constructing explicitly such a collection.

2. The construction

A q-ary *code* of length p, size q^m and distance d is a set of q^m vectors of length p over a set of q symbols $\{1, 2, \ldots, q\}$ such that any two vectors differ in at least

* Research supported in part by the Weizmann Fellowship for Scientific Research.

d coordinates, (i.e., their Hamming distance is $\geq d$). In [3, Theorem 5.7] Friedman uses the ideas of Justesen codes to explicitly construct, for every integer $l \geq 4$, every prime power q, $\frac{15}{2}l^2 \leq q \leq 15l^2$ and every integer $m \geq 1$ a code $C(l, q, m)$ of length $p = (15l^2)^{8l} \cdot m$, size q^m and distance $(1 - 3/l)p$. (This is, in fact, a somewhat relaxed version of Theorem 5.7 of [3]). Given $k \geq 2$, let $l = 3\binom{k}{2} + 1$ and let $v_1, v_2, \ldots, v_k$ be any k distinct vectors of $C(l, q, m)$, where $\frac{15}{2}l^2 \leq q \leq 15l^2$ is a prime power and $m \geq 1$. One can easily check that $v_1, v_2, \ldots, v_k$ attain k distinct values in at least one coordinate. Indeed, if this is false, then for every coordinate some pair of the v_i's agree and hence the sum of distances between the $\binom{k}{2}$ pairs of v_i's is $\leq p\binom{k}{2} - p = \binom{k}{2}p(1 - 1/\binom{k}{2}) < \binom{k}{2}(1 - 3/l)p$, contradicting the definition of our code. Let B be a binary matrix of 2^q rows and q columns, the ith row being the binary representation of $i - 1$ $(1 \leq i \leq 2^q)$. Note that for any choice of k distinct columns of B and any choice of k bits there are exactly 2^{q-k} rows of B that have the jth chosen bit in the jth chosen column. Let D be the set of q^m binary vectors of length $2^q p$ defined as follows: the ith vector u of D is obtained from the ith vector v of $C(l, q, m)$ by replacing any occurrence of j in v by the jth column of B $(1 \leq j \leq q)$. The discussion above clearly implies that for every k distinct vectors $u_1, \ldots, u_k$ of D and for any choice of k bits $b_1, b_2, \ldots, b_k$ there are at least 2^{q-k} coordinates in which u_i attains the value b_i $(1 \leq i \leq k)$. Hence, if we put $n = 2^q p$ and let F be the collection of subsets of $N = \{1, 2, \ldots, n\}$ whose characteristic vectors are the vectors in D we obtain an explicit k-independent collection of size q^m. Since $7l^2 \leq q \leq 15l^2$, $3k^2 \geq l = 3\binom{k}{2} + 1 \geq k^2$, $n = 2^q p$ and $p = (15l^2)^{8l} \cdot m$ one can easily check that

$$|F| \geq 2^{c_k \cdot n},$$

where

$$c_k = \log(7k^4)/(2^{135k^4} \cdot (135k^4)^{24k^2}) > 0. \tag{2.1}$$

This settles the problem raised in [4]. Note that the value of our c_k is much smaller than the one obtained by the non-constructive method of [4].

Let F be a 3-independent collection on $\{1, 2, \ldots, n\}$ of size $|F| \geq 2^{c_3 \cdot n}$ constructed above.

Let G be a collection of n-element subsets of $\{1, 2, \ldots, 2n\}$ of size $|F|$ obtained by adding to each $A \in F$ the first $n - |A|$ elements in $\{n + 1, \ldots, 2n\}$. Clearly the intersection of no two members of G is included in a third. This solves the problem raised in [5].

3. Remarks

(1) Our method easily supplies explicit examples of families F_r of subsets of $\{1, 2, \ldots, n\}$ in which no set is covered by the union of r others, where $|F_r| > 2^{d_r \cdot n}$ for some $d_r > 0$. The existence of such families was proved, non-constructively, by Erdös, Frankl and Füredi in [1, 2].

(2) We made no attempt to maximize the value of c_k given in (2.1) in our basic construction. One can easily improve (2.1) by a slightly more careful construction, but even the improved bound will be, of course, much smaller than the one supplied by the non-constructive method.

Acknowledgment

I am thankful to R. Boppana for showing me Friedman's paper and for fruitful discussions.

Note added in proof

Poljak, Pultr and Rödl [6, Theorem 3.7] gave explicit k-independent collections of almost exponential size. Their families are of size $2^{c_k n/(\log^{k-2} n)}$ (which is slightly weaker than our exponential bound), and they obtain similar explicit lower bounds for several, more general, problems. An obvious modification of our method here supplies exponential explicit lower bounds for these more general problems.

References

[1] P. Erdös, P. Frankl and Z. Füredi, Families of finite sets in which no set is covered by the union of two others, J. Combin. Theory Ser. A 33 (1982) 158–166.

[2] P. Erdös, P. Frankl and Z. Füredi, Families of finite sets in which no set is covered by the union of r others, preprint (1984).

[3] J. Friedman, Constructing $O(n \log n)$ size monotone formulae for the kth elementary symmetric polynomial of n boolean variables, Proc. 25th Annual Symp. on Foundations of Computer Science, IEEE, Florida (1984) 506–515.

[4] D.J. Kleitman and J. Spencer, Families of k-independent sets, Discrete Math. 6 (1973) 255–262.

[5] D.J. Kleitman, J. Shearer and D. Sturtevant, Intersections of k-element sets, Combinatorica 1 (1981) 381–384.

[6] S. Poljak, A. Pultr and V. Rödl, On qualitatively independent partitions and related problems, Discrete Appl. Math. 6 (1983) 193–205.