

Irredundant Families of Subcubes

David Ellis

DPMMS, University of Cambridge

We work in $Q_n = \{0, 1\}^n$.

Definition

A d -dimensional subcube of $\{0, 1\}^n$ is a subset of $\{0, 1\}^n$ of the form

$$\{x \in \{0, 1\}^n : x_i = a_i \forall i \in T\}$$

where T is a set of $n - d$ coordinates, called the *fixed coordinates*, and the a_i 's are fixed elements of $\{0, 1\}$.

The other coordinates $S = [n] \setminus T$ are called the *moving coordinates*.

So there are $\binom{n}{n-d} 2^{n-d}$ d -subcubes of $\{0, 1\}^n$.

Definition

A family \mathcal{A} of d -subcubes of $\{0, 1\}^n$ is *irredundant* if none of its subcubes is contained in the union of the others, i.e. each subcube has a 'private' vertex, not contained in any of the others.

Basic question:

What is the maximum possible size $M(n, d)$ of an irredundant family of d -subcubes of $\{0, 1\}^n$?

Conjecture (Aharoni-Holzman, 1991)

If $d > n/2$, then an irredundant family of d -subcubes of $\{0, 1\}^n$ has size at most $\binom{n}{d}$.

In its full strength, this conjecture remains open.

Notice that when $n = 5$ and $d = 3$, we don't have uniqueness: there is 'another' irredundant family of size $\binom{5}{3} = 10$:

$$\begin{pmatrix} 1 & 0 & 1 \\ *, & *, & * \\ 0, & 0, \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0, & *, & * \\ *, & *, & * \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0, & 0, & * \\ *, & *, & * \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ *, & 0, & 0 \\ *, & *, & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 1 \\ *, & *, & 0 \\ 0, & 0, & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ *, & *, & * \\ *, & 1, & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1, & *, & *, & * \\ *, & *, & *, & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1, & 1, & *, & *, & * \\ *, & *, & *, & *, & * \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ *, & 1, & 1, & *, & * \\ *, & *, & *, & *, & * \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ *, & *, & 1, & 1, & * \\ *, & *, & *, & *, & * \end{pmatrix}$$

Lemma (Aharoni-Holzman, 1991)

For any $d \leq n$,

$$M(n, d) \leq \sum_{i=0}^{n-d} \binom{n}{i}.$$

Proof:

Let \mathcal{A} be an irredundant family of d -subcubes of $\{0, 1\}^n$.

The characteristic functions

$$\{\chi_C : C \in \mathcal{A}\}$$

are linearly independent over \mathbb{R} : suppose

$$\sum_{C \in \mathcal{A}} a_C \chi_C = 0.$$

For each $D \in \mathcal{A}$, choose a private vertex w_D ;

Evaluating the above on w_D gives $a_D = 0$.

The characteristic function of a subcube C can be considered as an element of $\mathbb{R}[x_1, \dots, x_n]$:

if $T_0 =$ fixed 0's of C , $T_1 =$ fixed 1's of C , then

$$\chi_C(x_1, \dots, x_n) = \prod_{i \in T_0} (1 - x_i) \prod_{i \in T_1} x_i.$$

The set of monomials $S = \{\prod_{i \in I} x_i : I \in [n]^{\leq n-d}\}$ is a basis for the vector subspace

$$\text{Span}\{\chi_C : C \text{ is a } d\text{-subcube of } \{0, 1\}^n\} \subset \mathbb{R}[x_1, \dots, x_n]$$

which therefore has dimension

$$\sum_{i=0}^{n-d} \binom{n}{i}.$$

It follows that

$$|\mathcal{A}| \leq \sum_{i=0}^{n-d} \binom{n}{i}.$$

Theorem (Meshulam, 1992)

For any $d \leq n$,

$$M(n, d) \leq \frac{2^n}{\sum_{i=0}^d \binom{n}{i}} \binom{n}{d}.$$

Corollary

If $d \geq 9n/10$, then

$$M(n, d) = \binom{n}{d}.$$

An 'alternative' proof of Meshulam's Theorem:

Let \mathcal{A} be an irredundant family of d -subcubes of $\{0, 1\}^n$.

For each $C \in \mathcal{A}$, choose a private vertex w_C .

Fix $x \in \{0, 1\}^n$; let \mathcal{A}_x be the subcubes in \mathcal{A} which contain x .

We will show:

$$\sum_{C \in \mathcal{A}_x} \frac{1}{\binom{|w_C \Delta x| + n - d}{n - d}} \leq 1$$

Intuitively, this is saying that the w_C 's can't be too closely packed around x .

By translation, we may assume that $x = (0, \dots, 0) = O$, so we must show:

$$\sum_{C \in \mathcal{A}_O} \frac{1}{\binom{|w_C| + n - d}{n - d}} \leq 1.$$

This is an immediate consequence of Bollobás' Inequality:

Theorem (Bollobás)

Let $a_1, \dots, a_N, b_1, \dots, b_N \subset [n]$ such that

$$a_i \cap b_j = \emptyset \text{ iff } i = j.$$

Then

$$\sum_{i=1}^N \frac{1}{\binom{|a_i|+|b_i|}{|a_i|}} \leq 1.$$

For each $C \in \mathcal{A}_O$, let f_C be the set of fixed coordinates of C ; then

$$w_C \in D \text{ iff } w_C \cap f_D = \emptyset,$$

so

$$w_C \cap f_D = \emptyset \text{ iff } C = D,$$

and our claim follows from Bollobás' Inequality.

Summing the inequality

$$\sum_{C \in \mathcal{A}_x} \frac{1}{\binom{|w_C \Delta x| + n - d}{n - d}} \leq 1$$

over all $x \in \{0, 1\}^n$ proves Meshulam's Theorem:
Swapping the order of summation,

$$\begin{aligned} 2^n &\geq \sum_{C \in \mathcal{A}} \sum_{x \in C} \frac{1}{\binom{|w_C \Delta x| + n - d}{n - d}} \\ &= |\mathcal{A}| \sum_{i=0}^d \frac{\binom{d}{i}}{\binom{i + n - d}{n - d}} \\ &= \frac{|\mathcal{A}|}{\binom{n}{d}} \sum_{i=0}^d \binom{n}{i} \end{aligned}$$

So

$$|\mathcal{A}| \leq \frac{2^n}{\sum_{i=0}^d \binom{n}{i}} \binom{n}{d}.$$

Notice that equality holds in Meshulam's Theorem whenever there exists a *perfect d -error-correcting code*—equivalently, a partition of $\{0, 1\}^n$ into radius- d Hamming balls. But this doesn't happen very often...

Theorem (Tietäväinen, 1973)

There is a perfect d -error-correcting code in $\{0, 1\}^n$ precisely in the following cases:

- ▶ $d = 1$, $n + 1$ is a power of 2 (take any Hamming code)
- ▶ $d = 3$, $n = 23$ (take the Golay code)
- ▶ $n = 2d + 1$ (take a 'trivial' code, two vertices of distance n apart)

So

$$M(2d + 1, d) = 2 \binom{2d + 1}{d}.$$

In general, we can construct an irredundant family by:

- ▶ Taking a packing of Hamming balls of radius d in $\{0, 1\}^n$, or equivalently a $(2d + 1)$ -separated subset of $\{0, 1\}^n$;
- ▶ Taking all d -subcubes through the centre of each ball.

But this gives quite a small family when $d = \Omega(n)$...

For example, when $d = n/4$, it follows from a theorem of Corrádi and Katai that even an $(n/2)$ -separated family in $\{0, 1\}^n$ has size at most $2n$, so we get an irredundant family of size

$$\leq 2n \binom{n}{n/4} \leq 2n \exp\left(-\frac{n}{32}\right) 2^n$$

—exponentially smaller than the upper bound from Meshulam's theorem, which is

$$(2/3 + o(1))2^n.$$

However, it turns out that for $d = \lfloor \gamma n \rfloor$, where γ is fixed and $n \rightarrow \infty$, Meshulam's bound gives the right order of magnitude for $M(n, d)$...

We can achieve this with a probabilistic construction:

- ▶ Take a random set S where each vertex of $\{0, 1\}^n$ is present independently with probability p .
- ▶ Consider the (random) set W of all points in $\{0, 1\}^n$ of Hamming distance d from S .
- ▶ For each $w \in W$, choose any $x_w \in S$ of distance d from w , and let

$$C_w = \{y \in \{0, 1\}^n : y \Delta w \subset x_w \Delta w\}.$$

- ▶ Take the random family

$$\mathcal{A} = \{C_w : w \in W\}.$$

We can achieve this with a probabilistic construction:

- ▶ Take a random set S where each vertex of $\{0, 1\}^n$ is present independently with probability p .
- ▶ Consider the (random) set W of all points in $\{0, 1\}^n$ of Hamming distance d from S .
- ▶ For each $w \in W$, choose any $x_w \in S$ of distance d from w , and let

$$C_w = \{y \in \{0, 1\}^n : y \Delta w \subset x_w \Delta w\}.$$

- ▶ Take the random family

$$\mathcal{A} = \{C_w : w \in W\}.$$

The C_w 's are distinct;

(x_w is the unique point of S in C_w , and w is the 'opposite' point),
so $|\mathcal{A}| = |W|$.

Irredundant: w is a private vertex of C_w .

(If $w \in C_{w'}$, then we have $w, x_{w'} \in C_{w'}$, so $|w\Delta x_{w'}| \leq d$, so
 $|w\Delta x_{w'}| = d$, so $x_{w'} = x_w$.)

What is $\mathbb{E}|\mathcal{A}| = \mathbb{E}|W|$?

For $w \in \{0, 1\}^n$,

$$\begin{aligned}\mathbb{P}\{w \in W\} &= \mathbb{P}\{|w\Delta S| = d\} \\ &= (1 - p)^{\sum_{i=0}^{d-1} \binom{n}{i}} - (1 - p)^{\sum_{i=0}^d \binom{n}{i}}\end{aligned}$$

So

$$\mathbb{E}|\mathcal{A}| = 2^n \left((1 - p)^{\sum_{i=0}^{d-1} \binom{n}{i}} - (1 - p)^{\sum_{i=0}^d \binom{n}{i}} \right).$$

Let

$$\beta = \frac{\binom{n}{d}}{\sum_{i=0}^d \binom{n}{i}}, \quad y = (1 - p)^{\sum_{i=0}^d \binom{n}{i}};$$

then

$$\mathbb{E}|\mathcal{A}| = 2^n(y^{1-\beta} - y).$$

This attains its maximum of

$$\beta(1 - \beta)^{(1-\beta)/\beta}$$

when

$$y = (1 - \beta)^{1/\beta}.$$

Hence, choosing p such that

$$(1 - p)^{\sum_{i=0}^d \binom{n}{i}} = (1 - \beta)^{1/\beta},$$

our random irredundant family has expected size

$$\beta(1 - \beta)^{(1-\beta)/\beta} 2^n.$$

So

$$\beta(1-\beta)^{(1-\beta)/\beta}2^n \leq M(n, d) \leq \beta 2^n.$$

If $d = \lfloor \gamma n \rfloor$, where $\gamma < 1/2$ is fixed, then

$$\beta = \left(1 - \frac{\gamma}{1-\gamma}\right) (1 + o(1)),$$

so we get

$$2^n(1 + o(1)) \left(1 - \frac{\gamma}{1-\gamma}\right) \left(\frac{\gamma}{1-\gamma}\right)^{\frac{\gamma}{1-2\gamma}} \leq M(n, \lfloor \gamma n \rfloor) \leq 2^n(1 + o(1)) \left(1 - \frac{\gamma}{1-\gamma}\right).$$

If $d \geq (1/2 + \epsilon)n$ for $\epsilon > 0$ fixed, we have

$$\binom{n}{d} \leq M(n, d) \leq (1 + o_\epsilon(1)) \binom{n}{d}.$$

When $d \geq n/2$, it is natural to try to construct a large irredundant family of d -subcubes by taking some through $\mathbf{0} = (0, \dots, 0)$ and some through $\mathbf{1} = (1, \dots, 1)$. However, this approach is doomed to failure, even when $d = n/2$...

Theorem (E, 2007)

If \mathcal{A} is an irredundant family of d -subcubes of $\{0, 1\}^{2d}$ through $\mathbf{0}$ or $\mathbf{1}$, then $|\mathcal{A}| \leq \binom{2d}{d}$.