Bad words arising from generalized Fibonacci cubes

Sandi Klavžar
University of Ljubljana, Slovenia
University of Maribor, Slovenia
joint work with *Aleksandar Ilić, Yoomi Rho, Sergey Shpectorov*

Univ. Bordeaux, LaBRI, UMR5800, F-33400 Talence
CNRS, LaBRI, UMR5800, F-33400 Talence
April 26, 2013

Fibonacci cubes

Hypercubes

Definition

The *d*-cube Q_d or hypercube of dimension d:

Hypercubes

Definition

The *d*-cube Q_d or hypercube of dimension d:

- $V(Q_d) = \{b_1 b_2 \ldots b_d \mid b_i \in \{0,1\}\}$.

Hypercubes

Definition

The *d-cube* Q_d or *hypercube of dimension* d:

- $V(Q_d) = \{b_1 b_2 \ldots b_d \mid b_i \in \{0,1\}\}$.
- Two vertices are adjacent if the binary strings differ in one position.

Hypercubes

Definition

The d-cube Q_d or hypercube of dimension d:

- $V(Q_d) = \{b_1 b_2 \ldots b_d \mid b_i \in \{0, 1\}\}$.
- Two vertices are adjacent if the binary strings differ in one position.

- $|V(Q_d)| = 2^d$

Hypercubes

Definition

The *d*-cube Q_d or hypercube of dimension d:

- $V(Q_d) = \{b_1 b_2 \ldots b_d \mid b_i \in \{0, 1\}\}$.
- Two vertices are adjacent if the binary strings differ in one position.

- $|V(Q_d)| = 2^d$
- Q_d is d-regular

Hypercubes

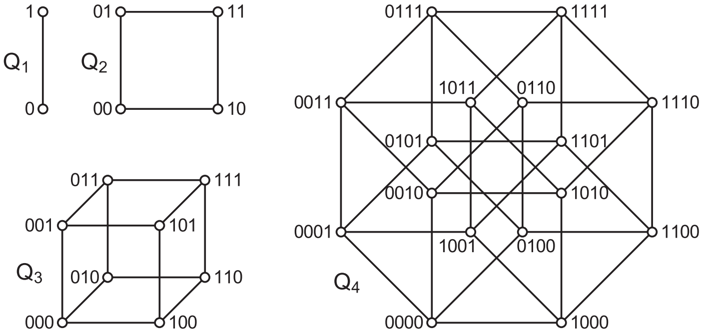
Definition

The *d-cube* Q_d or hypercube of dimension d:

- $V(Q_d) = \{b_1 b_2 \ldots b_d \mid b_i \in \{0, 1\}\}$.
- Two vertices are adjacent if the binary strings differ in one position.

- $|V(Q_d)| = 2^d$
- Q_d is d-regular
- Q_d is vertex-transitive ...

Small cubes


Q1
1
0
Q2
01
11
00
10
Q3
011
111
001
101
010
110
000
100
Q4
0111
1111
0011
1011
0110
1110
0101
1101
0010
1010
0001
1001
0100
1100
0000
1000

Defining Fibonacci cubes

- A Fibonacci string of length d is a binary string $b_1 b_2 \ldots b_d$ with $b_i \cdot b_{i+1} = 0$ for $1 \leq i < d$.

Defining Fibonacci cubes

- A Fibonacci string of length d is a binary string $b_1 b_2 \ldots b_d$ with $b_i \cdot b_{i+1} = 0$ for $1 \leq i < d$.
    - 01010001 and 00100001 are Fibonacci strings.
    - 011000 and 001010110 are not Fibonacci strings.

Defining Fibonacci cubes

- A Fibonacci string of length d is a binary string $b_1 b_2 \ldots b_d$ with $b_i \cdot b_{i+1} = 0$ for $1 \leq i < d$.
    - 01010001 and 00100001 are Fibonacci strings.
    - 011000 and 001010110 are not Fibonacci strings.
- Name comes from Zeckendorf's theorem.

Defining Fibonacci cubes

- A Fibonacci string of length d is a binary string $b_1 b_2 \ldots b_d$ with $b_i \cdot b_{i+1} = 0$ for $1 \leq i < d$.
  - 01010001 and 00100001 are Fibonacci strings.
  - 011000 and 001010110 are not Fibonacci strings.
- Name comes from Zeckendorf's theorem.

Definition

Fibonacci cube Γ_d, $d \geq 1$: subgraph of Q_d induced by the Fibonacci strings of length d.

Defining Fibonacci cubes

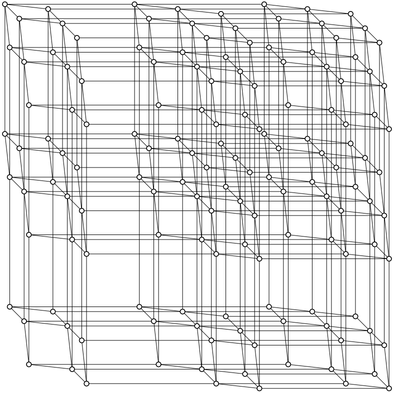
- A Fibonacci string of length d is a binary string $b_1 b_2 \ldots b_d$ with $b_i \cdot b_{i+1} = 0$ for $1 \leq i < d$.
  - 01010001 and 00100001 are Fibonacci strings.
  - 011000 and 001010110 are not Fibonacci strings.
- Name comes from Zeckendorf's theorem.

Definition

Fibonacci cube Γ_d, $d \geq 1$: subgraph of Q_d induced by the Fibonacci strings of length d.

Introduced in (Hsu, 1993) as a model for interconnection networks.

The Fibonacci cube Γ_{10}



Fibonacci cubes as simplex graphs

G graph, then the simplex graph $\kappa(G)$ of G:

Fibonacci cubes as simplex graphs

G graph, then the simplex graph $\kappa(G)$ of G:

- vertices: complete subgraphs of G (including empty),
- edges: complete subgraphs adjacent if differ in a single vertex.

Fibonacci cubes as simplex graphs

G graph, then the simplex graph $\kappa(G)$ of G:

- vertices: complete subgraphs of G (including empty),
- edges: complete subgraphs adjacent if differ in a single vertex.

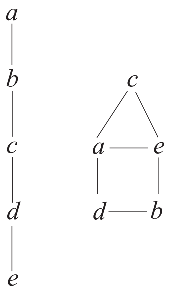
Γ_d is the simplex graph of the complement of P_d:

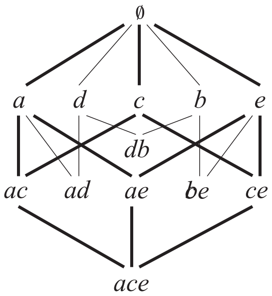
Fibonacci cubes as simplex graphs

G graph, then the simplex graph $\kappa(G)$ of G:

- vertices: complete subgraphs of G (including empty),
- edges: complete subgraphs adjacent if differ in a single vertex.

Γ_d is the simplex graph of the complement of P_d:





Selected results on F_d

Theorem

(Cong, Zheng, Sharma, 1993) Γ_d has a hamiltonian path for any $d \geq 0$.

Selected results on F_d

Theorem

(Cong, Zheng, Sharma, 1993) Γ_d *has a hamiltonian path for any* $d \geq 0$.

(Munarini, Perelli Cippo, Zagaglia Salvi, 2001):

$$|E(\Gamma_n)| = \frac{nF_{n+1} + 2(n+1)F_n}{5}\,.$$

Selected results on F_d

Theorem

(Cong, Zheng, Sharma, 1993) Γ_d has a hamiltonian path for any $d \geq 0$.

(Munarini, Perelli Cippo, Zagaglia Salvi, 2001):

$$|E(\Gamma_n)| = \frac{nF_{n+1} + 2(n+1)F_n}{5}\,.$$

The number of 4-cycles of Γ_n (K., 2005):

$$-\frac{3n}{25}F_{n+1} + \left(\frac{n^2}{10} + \frac{3n}{50} - \frac{1}{25}\right)F_n\,.$$

Selected results on F_d cont'd

Theorem

(K., Mollard, Petkovšek, 2011) The number of vertices of Γ_n *having degree* k *is* $\sum_{i=0}^{k} \binom{n-2i}{k-i}\binom{i+1}{n-k-i+1}$.

Selected results on F_d cont'd

Theorem

(K., Mollard, Petkovšek, 2011) The number of vertices of Γ_n having degree k is $\sum_{i=0}^{k} \binom{n-2i}{k-i}\binom{i+1}{n-k-i+1}$.

Theorem

(Castro, Mollard, 2012) Let $n \geq k \geq 1$, then the number of vertices of Γ_n with eccentricity k is $\binom{k}{n-k} + \binom{k-1}{n-k}$.

Selected results on F_d cont'd

Theorem

(K., Mollard, Petkovšek, 2011) The number of vertices of Γ_n having degree k is $\sum_{i=0}^{k} \binom{n-2i}{k-i}\binom{i+1}{n-k-i+1}$.

Theorem

(Castro, Mollard, 2012) Let $n \geq k \geq 1$, then the number of vertices of Γ_n with eccentricity k is $\binom{k}{n-k} + \binom{k-1}{n-k}$.

Theorem

(K., Mollard, 2012) $W(\Gamma_n) = \frac{4(n+1)F_n^2}{25} + \frac{(9n+2)F_nF_{n+1}}{25} + \frac{6nF_{n+1}^2}{25}$.

Selected results on F_d cont'd

Theorem

(Taranenko, Vesel, 2007) Fibonacci cubes can be recognized in $O(m \log n)$ time.

Selected results on F_d cont'd

Theorem

(Taranenko, Vesel, 2007) Fibonacci cubes can be recognized in $O(m \log n)$ time.

Theorem

(Vesel, 2013/14) Fibonacci cubes can be recognized in $O(m)$ time.

Generalized Fibonacci cubes

Defining generalized Fibonacci cubes

- f arbitrary binary string, $d \geq 1$.

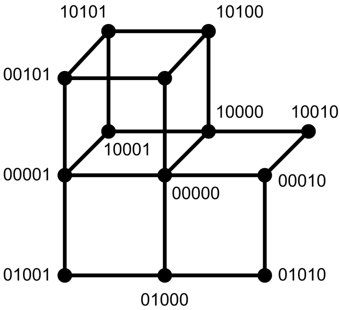
Defining generalized Fibonacci cubes

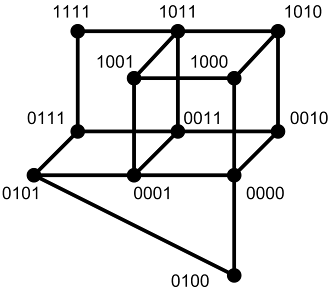
- f arbitrary binary string, $d \geq 1$.
- The generalized Fibonacci cube, $Q_d(f)$, is the graph obtained from Q_d by removing all vertices that contain f as a substring.

Defining generalized Fibonacci cubes

- f arbitrary binary string, $d \geq 1$.
- The generalized Fibonacci cube, $Q_d(f)$, is the graph obtained from Q_d by removing all vertices that contain f as a substring.
- Note: $\Gamma_d = Q_d(11)$.

Fibonacci cube $Q_5(11)$ and 110-Fibonacci cube $Q_4(110)$





A couple of properties

Set $H_d = Q_d(110)$.

A couple of properties

Set $H_d = Q_d(110)$.

Proposition

For any $d \geq 0$, $|E(H_d)| = -1 + \sum_{i=1}^{d+1} F_i F_{d+2-i}$.

A couple of properties

Set $H_d = Q_d(110)$.

Proposition

For any $d \geq 0$, $|E(H_d)| = -1 + \sum_{i=1}^{d+1} F_i F_{d+2-i}$.

Proposition

For any $d \geq 0$,

$$|S(H_d)| = -\frac{3(d+1)}{25} F_{d+2} + \left(\frac{(d+1)^2}{10} + \frac{3(d+1)}{50} - \frac{1}{25} \right) F_{d+1}.$$

Isometric embeddability

Definition

H is an isometric subgraph of G if for any $u, v \in V(H)$, $d_H(u, v) = d_G(u, v)$.

Isometric embeddability

Definition

H is an isometric subgraph of G if for any $u, v \in V(H)$, $d_H(u, v) = d_G(u, v)$.

- $H \hookrightarrow G$: H is an isometric subgraph of G
- $H \not\hookrightarrow G$: H is not an isometric subgraph of G

Isometric embeddability

Definition

H is an isometric subgraph of G if for any $u, v \in V(H)$, $d_H(u, v) = d_G(u, v)$.

- $H \hookrightarrow G$: H is an isometric subgraph of G
- $H \not\hookrightarrow G$: H is not an isometric subgraph of G

Problem

For which f and d, $Q_d(f) \hookrightarrow Q_d$ holds?

Isometric embeddability cont'd

- $\overline{b}$: binary complement of b,

Isometric embeddability cont'd

- $\overline{b}$: binary complement of b,
- b^R: reverse of b.

Isometric embeddability cont'd

- $\overline{b}$: binary complement of b,
- b^R: reverse of b.

Facts:

- f of length r, $1 \leq d \leq r$. Then $Q_d(f) \hookrightarrow Q_d$.

Isometric embeddability cont'd

- $\overline{b}$: binary complement of b,
- b^R: reverse of b.

Facts:

- f of length r, $1 \leq d \leq r$. Then $Q_d(f) \hookrightarrow Q_d$.
- $Q_d(f)$ is isomorphic to $Q_d(\overline{f})$.

Isometric embeddability cont'd

- $\overline{b}$: binary complement of b,
- b^R: reverse of b.

Facts:

- f of length r, $1 \leq d \leq r$. Then $Q_d(f) \hookrightarrow Q_d$.
- $Q_d(f)$ is isomorphic to $Q_d(\overline{f})$.
- $Q_d(f)$ is isomorphic to $Q_d(f^R)$.

Forbidden factors with at most three blocks

Proposition

Let $s \geq 1$. *Then* $Q_d(1^s) \hookrightarrow Q_d$.

Forbidden factors with at most three blocks

Proposition

Let $s \geq 1$*. Then* $Q_d(1^s) \hookrightarrow Q_d$.

Proposition

Let $r, s, t \geq 1$ *and let* $d \geq r + s + t + 1$*. Then* $Q_d(1^r 0^s 1^t) \not\hookrightarrow Q_d$.

Forbidden factors with at most three blocks

Proposition

Let $s \geq 1$*. Then* $Q_d(1^s) \hookrightarrow Q_d$.

Proposition

Let $r, s, t \geq 1$ *and let* $d \geq r + s + t + 1$*. Then* $Q_d(1^r 0^s 1^t) \not\hookrightarrow Q_d$.

Theorem

Let $d \geq 2$*. Then*

(i) *For* $r \geq 1$, $Q_d(1^r 0) \hookrightarrow Q_d$.

(ii) *For* $s \geq 2$, $Q_d(1^2 0^s) \hookrightarrow Q_d$ *if and only if* $d \leq s + 4$.

(iii) *If* $r, s \geq 3$, *then* $Q_d(1^r 0^s) \hookrightarrow Q_d$ *if and only if* $d \leq 2r + 2s - 3$.

Forbidden factors with more than three blocks

Proposition

Let $s \geq 1$*. Then* $Q_d((10)^s 1) \not\hookrightarrow Q_d$ *for* $d \geq 4s$*.*

Forbidden factors with more than three blocks

Proposition

Let $s \geq 1$. *Then* $Q_d((10)^s 1) \not\hookrightarrow Q_d$ *for* $d \geq 4s$.

Proposition

Let $r, s \geq 1$. *Then* $Q_d((10)^r 1 (10)^s) \not\hookrightarrow Q_d$ *for* $d \geq 2r + 2s + 3$.

Forbidden factors with more than three blocks

Proposition

Let $s \geq 1$*. Then* $Q_d((10)^s 1) \not\hookrightarrow Q_d$ *for* $d \geq 4s$*.*

Proposition

Let $r, s \geq 1$*. Then* $Q_d((10)^r 1 (10)^s) \not\hookrightarrow Q_d$ *for* $d \geq 2r + 2s + 3$*.*

Theorem

Let $s \geq 2$*. Then* $Q_d(1^s 0 1^s 0) \hookrightarrow Q_d$*.*

Forbidden factors with more than three blocks

Proposition

Let $s \geq 1$. *Then* $Q_d((10)^s 1) \not\hookrightarrow Q_d$ *for* $d \geq 4s$.

Proposition

Let $r, s \geq 1$. *Then* $Q_d((10)^r 1 (10)^s) \not\hookrightarrow Q_d$ *for* $d \geq 2r + 2s + 3$.

Theorem

Let $s \geq 2$. *Then* $Q_d(1^s 0 1^s 0) \hookrightarrow Q_d$.

Theorem

Let $s \geq 1$. *Then* $Q_d((10)^s) \hookrightarrow Q_d$.

Classification of strings of length at most five

- **11**, **10**,

Classification of strings of length at most five

- **11**, **10**,
- **111**, **110**, 101,

Classification of strings of length at most five

- **11**, **10**,
- **111**, **110**, 101,
- **1111**, **1110**, **1010**, 1101, 1001,

Classification of strings of length at most five

- **11**, **10**,
- **111**, **110**, 101,
- **1111**, **1110**, **1010**, 1101, 1001,
- **1100** ($d \leq 6$), 1100 ($d \geq 7$),

Classification of strings of length at most five

- **11**, **10**,
- **111**, **110**, 101,
- **1111**, **1110**, **1010**, 1101, 1001,
- **1100** ($d \leq 6$), 1100 ($d \geq 7$),
- **11111**, **11110**, **11010**, 11001, 11101, 11011, 10001,

Classification of strings of length at most five

- **11**, **10**,
- **111**, **110**, 101,
- **1111**, **1110**, **1010**, 1101, 1001,
- **1100** ($d \leq 6$), 1100 ($d \geq 7$),
- **11111**, **11110**, **11010**, 11001, 11101, 11011, 10001,
- **11100** ($d \leq 7$), 11100 ($d \geq 8$),

Classification of strings of length at most five

- **11**, **10**,
- **111**, **110**, 101,
- **1111**, **1110**, **1010**, 1101, 1001,
- **1100** ($d \leq 6$), 1100 ($d \geq 7$),
- **11111**, **11110**, **11010**, 11001, 11101, 11011, 10001,
- **11100** ($d \leq 7$), 11100 ($d \geq 8$),
- **10110** ($d \leq 6$), 10110 ($d \geq 7$),

Classification of strings of length at most five

- **11**, **10**,
- **111**, **110**, 101,
- **1111**, **1110**, **1010**, 1101, 1001,
- **1100** ($d \leq 6$), 1100 ($d \geq 7$),
- **11111**, **11110**, **11010**, 11001, 11101, 11011, 10001,
- **11100** ($d \leq 7$), 11100 ($d \geq 8$),
- **10110** ($d \leq 6$), 10110 ($d \geq 7$),
- **10101** ($d \leq 7$), 10101 ($d \geq 8$).

1: Good and bad words

Bad words

Definition

A word f is bad if there exists a dimension d such that $Q_d(f) \not\hookrightarrow Q_d$.

Bad words

Definition

A word f is bad if there exists a dimension d such that $Q_d(f) \not\hookrightarrow Q_d$.

Examples: $(10)^s1$ ($s \geq 1$), and 1^r0^s ($r, s \geq 2$), are bad words.

Bad words

Definition

A word f is bad if there exists a dimension d such that $Q_d(f) \not\hookrightarrow Q_d$.

Examples: $(10)^s1$ ($s \geq 1$), and 1^r0^s ($r, s \geq 2$), are bad words.

Lemma

Suppose that $Q_d(f) \not\hookrightarrow Q_d$ for some dimension d. Then $Q_{d'}(f) \not\hookrightarrow Q_{d'}$ for all $d' \geq d$.

Bad words cont'd

Proof

- $d' = d + r,\ r \geq 1.$

Bad words cont'd

Proof

- $d' = d + r$, $r \geq 1$.
- $Q_d(f) \not\hookrightarrow Q_d$ hence for some u and v, $d_{Q_d(f)}(u, v) > d_{Q_d}(u, v)$.

Bad words cont'd

Proof

- $d' = d + r$, $r \geq 1$.
- $Q_d(f) \not\hookrightarrow Q_d$ hence for some u and v, $d_{Q_d(f)}(u,v) > d_{Q_d}(u,v)$.
- If the first bit of f is 1, set $\widehat{u} = 0^r u$ and $\widehat{v} = 0^r v$, otherwise set $\widehat{u} = 1^r u$ and $\widehat{v} = 1^r v$.

Bad words cont'd

Proof

- $d' = d + r$, $r \geq 1$.
- $Q_d(f) \not\hookrightarrow Q_d$ hence for some u and v, $d_{Q_d(f)}(u,v) > d_{Q_d}(u,v)$.
- If the first bit of f is 1, set $\widehat{u} = 0^r u$ and $\widehat{v} = 0^r v$, otherwise set $\widehat{u} = 1^r u$ and $\widehat{v} = 1^r v$.
- $\widehat{u}, \widehat{v} \in Q_{d'}(f)$ and

Bad words cont'd

Proof

- $d' = d + r$, $r \geq 1$.
- $Q_d(f) \not\hookrightarrow Q_d$ hence for some u and v, $d_{Q_d(f)}(u,v) > d_{Q_d}(u,v)$.
- If the first bit of f is 1, set $\widehat{u} = 0^r u$ and $\widehat{v} = 0^r v$, otherwise set $\widehat{u} = 1^r u$ and $\widehat{v} = 1^r v$.
- $\widehat{u}, \widehat{v} \in Q_{d'}(f)$ and

$$d_{Q_{d'}(f)}(u,v) = d_{Q_d(f)}(u,v) > d_{Q_d}(u,v) = d_{Q_{d'}}(u,v)\,,$$

Bad words cont'd

Proof

- $d' = d + r$, $r \geq 1$.
- $Q_d(f) \not\hookrightarrow Q_d$ hence for some u and v, $d_{Q_d(f)}(u,v) > d_{Q_d}(u,v)$.
- If the first bit of f is 1, set $\widehat{u} = 0^r u$ and $\widehat{v} = 0^r v$, otherwise set $\widehat{u} = 1^r u$ and $\widehat{v} = 1^r v$.
- $\widehat{u}, \widehat{v} \in Q_{d'}(f)$ and

$$d_{Q_{d'}(f)}(u,v) = d_{Q_d(f)}(u,v) > d_{Q_d}(u,v) = d_{Q_{d'}}(u,v)\,,$$

hence $Q_{d'}(f) \not\hookrightarrow Q_{d'}$.

Good words

Definition

A word f is good if it is not bad.

Good words

Definition

A word f is good if it is not bad.

By the lemma, f is good if $Q_d(f) \hookrightarrow Q_d$ for all dimensions d.

Good words

Definition

A word f is good if it is not bad.

By the lemma, f is good if $Q_d(f) \hookrightarrow Q_d$ for all dimensions d.
Examples: 1^s, 10^s, and $(10)^s$, are good words for any $s \geq 1$.

Good words

Definition

A word f is good if it is not bad.

By the lemma, f is good if $Q_d(f) \hookrightarrow Q_d$ for all dimensions d.
Examples: 1^s, 10^s, and $(10)^s$, are good words for any $s \geq 1$.

$$\mathcal{G}_n = \{f \in B^n \mid f \text{ is good}\}$$

$$\mathcal{B}_n = \{f \in B^n \mid f \text{ is bad}\}$$

Combinatorics over words

- For a word f we introduce an r-error overlap of length k.

Combinatorics over words

- For a word f we introduce an r-error overlap of length k.
- If f has an r-error overlap for some length k then f has an r-error overlap.

Combinatorics over words

- For a word f we introduce an r-error overlap of length k.
- If f has an r-error overlap for some length k then f has an r-error overlap.
- $\mathcal{T}_n$ ... words of length n having a 2-error overlap.

- For a word f we introduce an r-error overlap of length k.
- If f has an r-error overlap for some length k then f has an r-error overlap.
- $\mathcal{T}_n$ ... words of length n having a 2-error overlap.
- $f \in \mathcal{T}_n$ is split if f has a 2-error overlap of length $k \leq \frac{n}{2}$.

- For a word f we introduce an r-error overlap of length k.
- If f has an r-error overlap for some length k then f has an r-error overlap.
- $\mathcal{T}_n$ ... words of length n having a 2-error overlap.
- $f \in \mathcal{T}_n$ is split if f has a 2-error overlap of length $k \leq \frac{n}{2}$.
- $\mathcal{T}_n^s$ ... split words from $\mathcal{T}_n$.

Lemma

The sequence $\frac{|\mathcal{T}_n^s|}{2^n}$ *is monotonically increasing and bounded from above by* 1*. In particular, it has a limit* $a \leq 1$*.*

Lemma

The sequence $\frac{|\mathcal{T}_n^s|}{2^n}$ *is monotonically increasing and bounded from above by* 1*. In particular, it has a limit* $a \leq 1$*.*

Lemma

The sequence $\frac{|\mathcal{T}_n|}{2^n}$ *converges to the same limit value* a*.*

Lemma

The sequence $\frac{|\mathcal{T}_n^s|}{2^n}$ *is monotonically increasing and bounded from above by* 1*. In particular, it has a limit* $a \leq 1$*.*

Lemma

The sequence $\frac{|\mathcal{T}_n|}{2^n}$ *converges to the same limit value* a*.*

Lemma

If f *is bad then* f *has a 2-error overlap.*

Lemma

The sequence $\frac{|\mathcal{T}_n^s|}{2^n}$ *is monotonically increasing and bounded from above by* 1. *In particular, it has a limit* $a \leq 1$.

Lemma

The sequence $\frac{|\mathcal{T}_n|}{2^n}$ *converges to the same limit value* a.

Lemma

If f *is bad then* f *has a* 2*-error overlap.*

Theorem

$$\lim_{n\to\infty} \frac{|\mathcal{B}_n|}{2^n} = a\,.$$

Estimates for the limit value a

n	$\lvert\mathcal{T}_n^s\rvert$	$\lvert\mathcal{T}_n^s\rvert/2^n$
4	4	0.250000
6	34	0.531250
8	182	0.710938
10	830	0.810547
12	3518	0.858887
14	14538	0.887329
16	59074	0.901398
18	238534	0.909935
22	3845886	0.916931
24	15408114	0.918395
26	61689006	0.919238
28	246881258	0.919704
30	987815218	0.919975

n	$\|\mathcal{T}_n\|$	$\|\mathcal{T}_n\|/2^n$
4	8	0.500000
5	22	0.687500
6	46	0.718750
7	98	0.765625
8	210	0.820313
9	430	0.839844
10	886	0.865234
25	30873042	0.920088
26	61759618	0.920290
27	123512490	0.920240
28	247051278	0.920338
29	494077866	0.920292
30	988213906	0.920346
31	1976359510	0.920314

Estimate for the limit value

Theorem

The value of the limit a is between 0.919975 *and* 0.924156.

Estimate for the limit value

Theorem

The value of the limit a is between 0.919975 *and* 0.924156.

Proof

- Lower bound is easy since the sequence a_n (densities of split words with a 2-error overlap) is monotonically increasing.

Estimate for the limit value

Theorem

The value of the limit a is between 0.919975 *and* 0.924156*.*

Proof

- Lower bound is easy since the sequence a_n (densities of split words with a 2-error overlap) is monotonically increasing.
- $T_n \ldots \#$ of nonsplit words of length n.

Estimate for the limit value

Theorem

The value of the limit a is between 0.919975 *and* 0.924156.

Proof

- Lower bound is easy since the sequence a_n (densities of split words with a 2-error overlap) is monotonically increasing.
- T_n ... # of nonsplit words of length n.
- Focus on words of even length $n = 2k$.

Estimate for the limit value

Theorem

The value of the limit a is between 0.919975 *and* 0.924156*.*

Proof

- Lower bound is easy since the sequence a_n (densities of split words with a 2-error overlap) is monotonically increasing.
- T_n ... # of nonsplit words of length n.
- Focus on words of even length $n = 2k$.
- $4T_n \leq T_{n+2} + 2^{k+1}\binom{k+1}{2}$.

Proof cont'd

- Let $\mu_n = T_n/2^n$ (notice $\mu_n = 1 - a_n$), then

$$\mu_n \leq \mu_{n+2} + \frac{k(k+1)}{2^{k+2}}\,.$$

Proof cont'd

- Let $\mu_n = T_n/2^n$ (notice $\mu_n = 1 - a_n$), then

$$\mu_n \le \mu_{n+2} + \frac{k(k+1)}{2^{k+2}} .$$

- This implies that

$$\mu_{n+2} \ge \mu_n - \frac{k(k+1)}{2^{k+2}} .$$

Proof cont'd

- Let $\mu_n = T_n/2^n$ (notice $\mu_n = 1 - a_n$), then

$$\mu_n \leq \mu_{n+2} + \frac{k(k+1)}{2^{k+2}}.$$

- This implies that

$$\mu_{n+2} \geq \mu_n - \frac{k(k+1)}{2^{k+2}}.$$

- Since $a_n = 1 - \mu_n$ we get

$$a_{n+2} \leq a_n + \frac{k(k+1)}{2^{k+2}}.$$

Proof cont'd

- Combining these relations from n to $n+2m$ we get

$$a_{n+2m} \leq a_n + \sum_{i=k}^{k+m-1} \frac{i(i+1)}{2^{i+2}} \, .$$

Proof cont'd

- Combining these relations from n to $n+2m$ we get

$$a_{n+2m} \leq a_n + \sum_{i=k}^{k+m-1} \frac{i(i+1)}{2^{i+2}}\,.$$

- Take m to infinity to get:

$$a \leq a_n + \sum_{i=k}^{\infty} \frac{i(i+1)}{2^{i+2}}\,.$$

Proof cont'd

- Combining these relations from n to $n + 2m$ we get

$$a_{n+2m} \leq a_n + \sum_{i=k}^{k+m-1} \frac{i(i+1)}{2^{i+2}} \,.$$

- Take m to infinity to get:

$$a \leq a_n + \sum_{i=k}^{\infty} \frac{i(i+1)}{2^{i+2}} \,.$$

- Using computer with $k = 15$ we get that the sum here is at most 0.004181.

Proof cont'd

- Combining these relations from n to $n + 2m$ we get

$$a_{n+2m} \leq a_n + \sum_{i=k}^{k+m-1} \frac{i(i+1)}{2^{i+2}} \,.$$

- Take m to infinity to get:

$$a \leq a_n + \sum_{i=k}^{\infty} \frac{i(i+1)}{2^{i+2}} \,.$$

- Using computer with $k = 15$ we get that the sum here is at most 0.004181.
- Together with the value $a_{30} = 0.919975$ this yields an upper limit of 0.924156.

2: Index of binary words

Index

Definition

The index of a word f, denoted $\beta(f)$, is the smallest integer d such that $Q_d(f) \not\hookrightarrow Q_d$. If no such integer exists we set $\beta(f) = \infty$.

Index

Definition

The index of a word f, denoted $\beta(f)$, is the smallest integer d such that $Q_d(f) \not\hookrightarrow Q_d$. If no such integer exists we set $\beta(f) = \infty$.

Clearly, $\beta(f) < \infty$ if and only if f is bad.

Index

Definition

The index of a word f, denoted $\beta(f)$, is the smallest integer d such that $Q_d(f) \not\hookrightarrow Q_d$. If no such integer exists we set $\beta(f) = \infty$.

Clearly, $\beta(f) < \infty$ if and only if f is bad.

Theorem

Let f be a bad word. Then $\beta(f) < |f|^2$.

Index cont'd

Theorem

For almost all bad words f, $\beta(f) < 2|f|$.

Index cont'd

Theorem

For almost all bad words f, $\beta(f) < 2|f|$.

Conjecture

For any bad word f, $\beta(f) < 2|f|$.

Index cont'd

Theorem

For almost all bad words f, $\beta(f) < 2|f|$.

Conjecture

For any bad word f, $\beta(f) < 2|f|$.

Conjecture verified by computer for all words of length at most 10 and dimension at most 20.

3: Parity index of binary words

Motivation - hamiltonicity

Recall: Γ_d has a hamiltonian path for any $d \geq 0$.

Recall: Γ_d has a hamiltonian path for any $d \geq 0$.

Theorem

(Liu, Hsu, Chung, 1994) Each $Q_d(1^r)$ contains a hamiltonian path.

Motivation - hamiltonicity

Recall: Γ_d has a hamiltonian path for any $d \geq 0$.

Theorem

(Liu, Hsu, Chung, 1994) Each $Q_d(1^r)$ contains a hamiltonian path.

Question: what about $Q_d(f)$?

Parity index defined

- Even/odd words of $Q_d(f)$ form its bipartition.

Parity index defined

- Even/odd words of $Q_d(f)$ form its bipartition.
- For a set of words X, let $e(X)$, $o(X)$ be the number of even/odd words in X.

Parity index defined

- Even/odd words of $Q_d(f)$ form its bipartition.
- For a set of words X, let $e(X)$, $o(X)$ be the number of even/odd words in X.
- $\Delta(X) = e(X) - o(X)$.

Parity index defined

- Even/odd words of $Q_d(f)$ form its bipartition.
- For a set of words X, let $e(X)$, $o(X)$ be the number of even/odd words in X.
- $\Delta(X) = e(X) - o(X)$.
- $\mathrm{PI}_d(f) = \Delta(\mathcal{F}_d(f))$ ... parity index of f of dimension d.

Parity index defined

- Even/odd words of $Q_d(f)$ form its bipartition.
- For a set of words X, let $e(X)$, $o(X)$ be the number of even/odd words in X.
- $\Delta(X) = e(X) - o(X)$.
- $\mathrm{PI}_d(f) = \Delta(\mathcal{F}_d(f))$ ... parity index of f of dimension d.

Therefore, a necessary condition for $Q_d(f)$ to contain a hamiltonian path is:

$$|\mathrm{PI}_d(f)| \leq 1\,.$$

Prime words

Definition

A word f of length d is prime if for any k, $1 \leq k \leq d-1$, the suffix of f of length k is different from the prefix of f of the same length.

Prime words

Definition

A word f of length d is prime if for any k, $1 \leq k \leq d-1$, the suffix of f of length k is different from the prefix of f of the same length.

001101 is a prime word (00 appears only at its beginning).

Prime words

Definition

A word f of length d is prime if for any k, $1 \leq k \leq d-1$, the suffix of f of length k is different from the prefix of f of the same length.

001101 is a prime word (00 appears only at its beginning).
01101011 is not prime (it starts and ends with 011).

Prime words

Definition

A word f of length d is prime if for any k, $1 \leq k \leq d-1$, the suffix of f of length k is different from the prefix of f of the same length.

001101 is a prime word (00 appears only at its beginning).
01101011 is not prime (it starts and ends with 011).

Definition

A word b is a power of a word c if $b = c^k$ for some $k \geq 1$.

Prime words

Definition

A word f of length d is prime if for any k, $1 \leq k \leq d-1$, the suffix of f of length k is different from the prefix of f of the same length.

001101 is a prime word (00 appears only at its beginning).
01101011 is not prime (it starts and ends with 011).

Definition

A word b is a power of a word c if $b = c^k$ for some $k \geq 1$.

Theorem

Let f be a power of a prime word. Then $|\mathrm{PI}_d(f)| \leq 1$ for any $d \geq 1$.

Prime words cont'd

Conjecture

Let f be a word such that $|\mathrm{PI}_d(f)| \leq 1$ holds for any d. Then f is a power of a prime word.

Prime words cont'd

Conjecture

Let f be a word such that $|\mathrm{PI}_d(f)| \le 1$ holds for any d. Then f is a power of a prime word.

Theorem

Let $r \ge 1$. Then

$$|\mathrm{PI}_d(0^r 1 0^r)| = \begin{cases} 0; & d \le 2r, 2r+2 \le d \le 3r+1\,, \\ 1; & d = 2r+1, 3r+2 \le d \le 4r+3\,. \end{cases}$$

Moreover, for any $d \ge 4r+4$, $|\mathrm{PI}_d(0^r 1 0^r)| \ge 2$.

Support for the conjecture

Theorem

Let $r, s, t \geq 1$. Let z be the integer such that $(z-1)t + 2 \leq r + s \leq zt + 1$. Then

$$|\mathrm{PI}_d(0^r 1^s 0^t)| \begin{cases} = 0; & d < r+s+t, \\ & y(r+s+t) < d < (y+1)(r+s)+t; 1 \leq y \leq z\,, \\ \geq 1; & d = r+s+t, \\ & (y+1)(r+s)+t \leq d \leq (y+1)(r+s+t); 1 \leq y \leq z, \\ & d = (z+1)(r+s+t)+1\,. \end{cases}$$

Moreover, for any $d \geq (z+1)(r+s+t)+2$, $|\mathrm{PI}_d(0^r 1^s 0^t)| \geq 2$.

Support for the conjecture cont'd

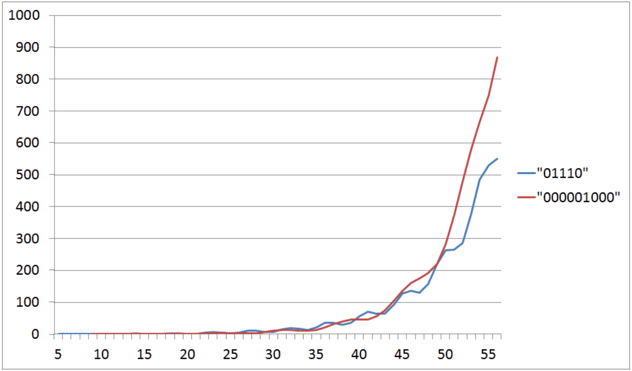
$\mathrm{PI}_d(f)$ computed for $|f| \leq 10$, $d \leq 31$. Balanced words:

Support for the conjecture cont'd

$\mathrm{PI}_d(f)$ computed for $|f| \leq 10$, $d \leq 31$. Balanced words:

length	f
3	001
4	0001, 0011, 0101
5	00001, 00011, 00101
6	000001, 000011, 000101, 000111 001001, 001011, 001101, 010101
7	0000001, 0000011, 0000101, 0000111 0001001, 0001011, 0001101, 0010011 0010101, 0011101
8	00000001, 00000011, 00000101, 00000111 00001001, 00001011, 00001101, 00001111 00010001, 00010011, 00010101, 00010111 00011001, 00011011, 00011101, 00100011 00100101, 00101011, 00101101, 00110011 00110101, 00111101, 01010101

Values of $|\mathrm{PI}_d(f)|$ for $f = 01110$ and $f = 000001000$



Merci beaucoup!