Thue choice number and the counting argument

Matthieu Rosenfeld

February 25, 2024

Square free words

A *square* is a non-empty word of the form uu. The period of uu is $|u|$

A word is *square-free* if none of its factor is a square.

Square free words

A *square* is a non-empty word of the form uu. The period of uu is $|u|$

A word is *square-free* if none of its factor is a square.

$abcabc$ is a square.

Square free words

A *square* is a non-empty word of the form uu. The period of uu is $|u|$

A word is *square-free* if none of its factor is a square.

$abcabc$ is a square.

$ba\underline{bcbc}abc$ is not square-free.

Square free words

A *square* is a non-empty word of the form uu. The period of uu is $|u|$

A word is *square-free* if none of its factor is a square.

$abcabc$ is a square.

$ba\underline{bcbc}abc$ is not square-free.

$abcacbac$ is square-free.

An old result

The starting point of combinatorics on words.

Theorem (Thue, 1906)

There exists arbitrarily long square-free words over $\{0, 1, 2\}$.

An old result

The starting point of combinatorics on words.

Theorem (Thue, 1906)

There exists arbitrarily long square-free words over $\{0, 1, 2\}$.

Many generalizations or variations were studied:

- Cubes, 4th powers, fractional powers,
- patterns, formulas (*ABABA*),
- k-abelian powers, k-binomial powers, additive powers, antipowers,
- nonrepetitive colorings of graphs (or other objects).
- ...

Thue choice number

Fix a sequence of alphabet $(A_i)_{i \in \mathbb{N}}$

A word $w = w_1 \dots w_n$ **respects** $(A_i)_{i \in \mathbb{N}}$ if for all i, $w_i \in A_i$

Thue choice number

Fix a sequence of alphabet $(A_i)_{i\in\mathbb{N}}$

A word $w = w_1 \dots w_n$ **respects** $(A_i)_{i\in\mathbb{N}}$ if for all i, $w_i \in A_i$

If w is infinite, we write $w \in \prod_i A_i$

Thue choice number

Fix a sequence of alphabet $(A_i)_{i \in \mathbb{N}}$

A word $w = w_1 \dots w_n$ **respects** $(A_i)_{i \in \mathbb{N}}$ if for all i, $w_i \in A_i$

If w is infinite, we write $w \in \prod_i A_i$

Alphabets: $\begin{Bmatrix} 0 \\ 1 \\ 2 \end{Bmatrix} \quad \begin{Bmatrix} 0 \\ 2 \\ 3 \end{Bmatrix} \quad \begin{Bmatrix} 1 \\ 2 \\ 3 \end{Bmatrix} \quad \begin{Bmatrix} 0 \\ 1 \\ 3 \end{Bmatrix} \quad \begin{Bmatrix} 0 \\ 2 \\ 3 \end{Bmatrix} \quad \begin{Bmatrix} 0 \\ 1 \\ 2 \end{Bmatrix} \quad \dots$

Thue choice number

Fix a sequence of alphabet $(A_i)_{i \in \mathbb{N}}$

A word $w = w_1 \dots w_n$ **respects** $(A_i)_{i \in \mathbb{N}}$ if for all i, $w_i \in A_i$

If w is infinite, we write $w \in \prod_i A_i$

Alphabets:	$\{0,1,2\}$	$\{0,2,3\}$	$\{1,2,3\}$	$\{0,1,3\}$	$\{0,2,3\}$	$\{0,1,2\}$	...
word:	0	2	1	3	0	2	...

Thue choice number

Fix a sequence of alphabet $(A_i)_{i \in \mathbb{N}}$

A word $w = w_1 \dots w_n$ **respects** $(A_i)_{i \in \mathbb{N}}$ if for all i, $w_i \in A_i$

If w is infinite, we write $w \in \prod_i A_i$

Alphabets:	$\{0, 1, 2\}$	$\{0, 2, 3\}$	$\{1, 2, 3\}$	$\{0, 1, 3\}$	$\{0, 2, 3\}$	$\{0, 1, 2\}$	...
word:	0	2	1	3	0	2	...

Question

Does there exists $k \in \mathbb{N}$, such that:

$$\forall i, |A_i| \geq k \implies \text{there exists a square-free word } w \in \prod_i A_i$$

Thue choice number

Fix a sequence of alphabet $(A_i)_{i\in\mathbb{N}}$

A word $w = w_1 \dots w_n$ **respects** $(A_i)_{i\in\mathbb{N}}$ if for all i, $w_i \in A_i$

If w is infinite, we write $w \in \prod_i A_i$

Alphabets:	$\{0,1,2\}$	$\{0,2,3\}$	$\{1,2,3\}$	$\{0,1,3\}$	$\{0,2,3\}$	$\{0,1,2\}$	...
word:	0	2	1	3	0	2	...

Question

Does there exists $k \in \mathbb{N}$, such that:

$$\forall i, |A_i| \geq k \implies \text{there exists a square-free word } w \in \prod_i A_i$$

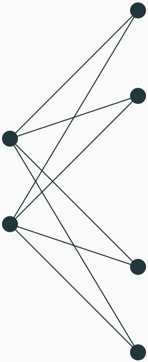
The Thue choice number is the smallest such k.

A short detour by graph coloring

A proper coloring of a graph G, is a coloring such that two adjacent vertices receive different colors

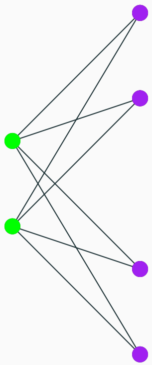
A short detour by graph coloring

A proper coloring of a graph G, is a coloring such that two adjacent vertices receive different colors



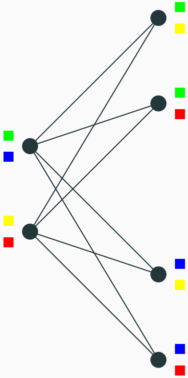
A short detour by graph coloring

A proper coloring of a graph G, is a coloring such that two adjacent vertices receive different colors



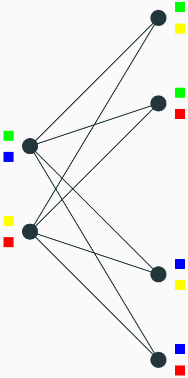
A short detour by graph coloring

A proper coloring of a graph G, is a coloring such that two adjacent vertices receive different colors



A short detour by graph coloring

A proper coloring of a graph G, is a coloring such that two adjacent vertices receive different colors



We have graphs with $\chi(G) = 2$ and $\chi_{ch}(G)$ arbitrarily large.

Avoiding squares over lists of size 4

Theorem (Grytczuk, Przybyło and Zhu (2011), ..., Rosenfeld (2020))

Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of lists such that for all i, $|A_i| \geq 4$. Then there are square-free words in $\prod_i A_i$.

Avoiding squares over lists of size 4

Theorem (Grytczuk, Przybyło and Zhu (2011), ..., Rosenfeld (2020))

Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of lists such that for all i, $|A_i| \geq 4$. Then there are square-free words in $\prod_i A_i$.

We show instead a stronger result.

Fix $(A_i)_{i\in\mathbb{N}}$. Let C_n be the set of square-free words of length n that respect $(A_i)_{i\in\mathbb{N}}$.

Lemma

For any integer n,

$$|C_{n+1}| \geq 2|C_n| \, .$$

Avoiding squares over lists of size 4

Theorem (Grytczuk, Przybyło and Zhu (2011), ..., Rosenfeld (2020))

Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of lists such that for all i, $|A_i| \geq 4$. Then there are square-free words in $\prod_i A_i$.

We show instead a stronger result.

Fix $(A_i)_{i\in\mathbb{N}}$. Let C_n be the set of square-free words of length n that respect $(A_i)_{i\in\mathbb{N}}$.

Lemma

For any integer n,

$$|C_{n+1}| \geq 2|C_n| \ .$$

$\implies |C_n| \geq 2^n \geq 1.$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$,

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}|$$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}|$$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i.

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,

$$\boxed{w \in \mathcal{B}_i}$$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

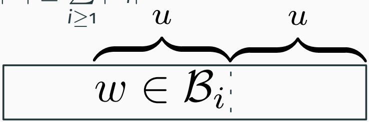
By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,



The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

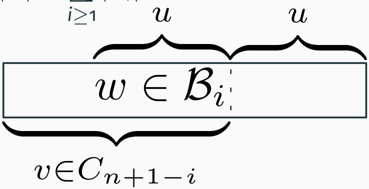
By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,



The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

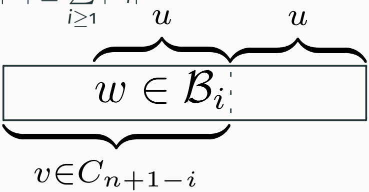
By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,



$|\mathcal{B}_i| \leq |C_{n+1-i}|$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \dots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

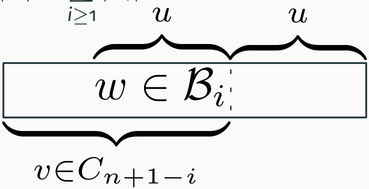
Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,

$$|\mathcal{B}_i| \leq |C_{n+1-i}| \leq \frac{|C_n|}{2^{i-1}} .$$



The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \dots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

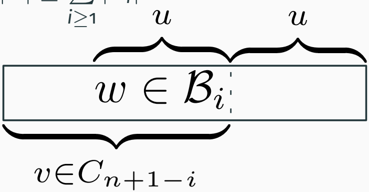
Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,

$$|\mathcal{B}_i| \leq |C_{n+1-i}| \leq \frac{|C_n|}{2^{i-1}} .$$



Finally,

$$|C_{n+1}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

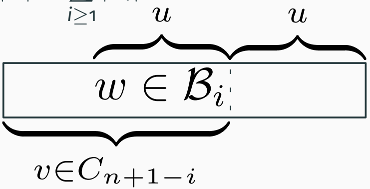
Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,

$|\mathcal{B}_i| \leq |C_{n+1-i}| \leq \frac{|C_n|}{2^{i-1}}$.



Finally,

$$|C_{n+1}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i| \geq 4|C_n| - \sum_{i \geq 1} \frac{|C_n|}{2^{i-1}}$$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

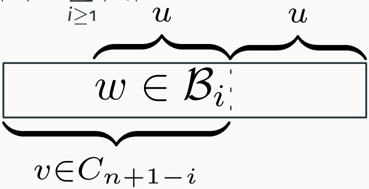
Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i \geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i \geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,

$$|\mathcal{B}_i| \leq |C_{n+1-i}| \leq \frac{|C_n|}{2^{i-1}} .$$



Finally,

$$|C_{n+1}| \geq 4|C_n| - \sum_{i \geq 1} |\mathcal{B}_i| \geq 4|C_n| - \sum_{i \geq 1} \frac{|C_n|}{2^{i-1}} = |C_n| \left(4 - \sum_{i \geq 0} 2^{-i} \right)$$

The proof by induction that for all n, $|C_{n+1}| \geq 2|C_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|C_{n-j}| \leq \frac{|C_n|}{2^j}$.

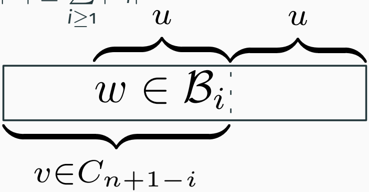
Let $\mathcal{B} = \{ua : u \in C_n, a \in A_{n+1}\} \setminus C_{n+1}$, then

$$|C_{n+1}| = |A_{n+1}| \cdot |C_n| - |\mathcal{B}| \geq 4|C_n| - |\mathcal{B}| \geq 4|C_n| - \sum_{i\geq 1} |\mathcal{B}_i|$$

For all i, let $\mathcal{B}_i$ be the set of words from $\mathcal{B}$ that end with a square of period i. Then $\mathcal{B} = \bigcup_{i\geq 1} \mathcal{B}_i$ and $|\mathcal{B}| \leq \sum_{i\geq 1} |\mathcal{B}_i|$.

For all $w \in \mathcal{B}_i$,

$$|\mathcal{B}_i| \leq |C_{n+1-i}| \leq \frac{|C_n|}{2^{i-1}} .$$



Finally,

$$|C_{n+1}| \geq 4|C_n| - \sum_{i\geq 1} |\mathcal{B}_i| \geq 4|C_n| - \sum_{i\geq 1} \frac{|C_n|}{2^{i-1}} = |C_n| \left(4 - \sum_{i\geq 0} 2^{-i} \right) \geq 2|C_n| \quad \square$$

The Thue choice number

Theorem (Grytczuk, Przybyło and Zhu (2011), ..., Rosenfeld (2020))

$$\forall i, |A_i| \geq 4 \implies \textit{there exists a square-free word } w \in \prod_i A_i$$

The Thue choice number

Theorem (Grytczuk, Przybyło and Zhu (2011), ..., Rosenfeld (2020))

$$\forall i, |A_i| \geq 4 \implies \textit{there exists a square-free word } w \in \prod_i A_i$$

Question

Is the following true ?

$$\forall i, |A_i| \geq 3 \implies \text{there exists a square-free word } w \in \prod_i A_i$$

Sharpening the proof - I

The main reason for which the proof is not sharp is this bound

$$|\mathcal{B}_i| \leq \frac{|C_n|}{2^{i-1}}$$

Sharpening the proof - I

The main reason for which the proof is not sharp is this bound

$$|\mathcal{B}_i| \leq \frac{|C_n|}{2^{i-1}}$$

For large i, it doesn't really matter since the bound is really small anyway.

Sharpening the proof - I

The main reason for which the proof is not sharp is this bound

$$|\mathcal{B}_i| \leq \frac{|C_n|}{2^{i-1}}$$

For large i, it doesn't really matter since the bound is really small anyway.

How to improve the bound for small i?

Compute the growth rate α_ℓ of the language of words without squares of period less than ℓ and hope for something like:

$$|C_{n+1}| \gtrsim \alpha_\ell |C_n| - \sum_{i \geq \ell} |\mathcal{B}_i|$$

Sharpening the proof sharp - II

Using elementary automata theory and linear algebra, we define a weight for each word and

$$\widehat{C_n} := \sum_{w \in C_n} \text{weight of } w$$

Sharpening the proof sharp - II

Using elementary automata theory and linear algebra, we define a weight for each word and

$$\widehat{C_n} := \sum_{w \in C_n} \text{weight of } w$$

such that

$$\widehat{C_{n+1}} \geq \alpha_\ell \widehat{C_n} - \sum_{i \geq \ell} \widehat{\mathcal{B}_i}$$

Sharpening the proof sharp - II

Using elementary automata theory and linear algebra, we define a weight for each word and

$$\widehat{C_n} := \sum_{w \in C_n} \text{weight of } w$$

such that

$$\widehat{C_{n+1}} \geq \alpha_\ell \widehat{C_n} - \sum_{i \geq \ell} \widehat{\mathcal{B}_i}$$

- α_ℓ is the spectral radius of the automaton,
- the weight function is given by the associated eigenvector,
- we can bound $\widehat{\mathcal{B}_i}$ with some bijection.

Sharpening the proof sharp - II

Using elementary automata theory and linear algebra, we define a weight for each word and

$$\widehat{C_n} := \sum_{w \in C_n} \text{weight of } w$$

such that

$$\widehat{C_{n+1}} \geq \alpha_\ell \widehat{C_n} - \sum_{i \geq \ell} \widehat{\mathcal{B}_i}$$

- α_ℓ is the spectral radius of the automaton,
- the weight function is given by the associated eigenvector,
- we can bound $\widehat{\mathcal{B}_i}$ with some bijection.

A similar idea was used by Kolpakov, Rao, Shur to obtain bounds on the growth of power-free languages

Back to the Thue choice number

Theorem (Rosenfeld, 2023)

Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of lists such that

- *for all i,* $|A_i| \geq 3$
- *and for all i,* $A_i \subseteq \{0, 1, 2, 3\}$.

Then there are square-free words in $\prod A_i$.

Back to the Thue choice number

Theorem (Rosenfeld, 2023)

Let $(A_i)_{i\in\mathbb{N}}$ *be a sequence of lists such that*

- *for all i,* $|A_i| \geq 3$
- *and for all i,* $A_i \subseteq \{0, 1, 2, 3\}$.

Then there are square-free words in $\prod A_i$.

Relies on the computation of the automaton recognizing words over $\{0, 1, 2, 3\}$ that do not contain squares of period less than 22.

(854683883 states, 70GB in memory, 2 hours of computation)

Back to the Thue choice number

Theorem (Rosenfeld, 2023)

Let $(A_i)_{i\in\mathbb{N}}$ *be a sequence of lists such that*

- *for all i,* $|A_i| \geq 3$
- *and for all i,* $A_i \subseteq \{0, 1, 2, 3\}$.

Then there are square-free words in $\prod A_i$.

Relies on the computation of the automaton recognizing words over $\{0, 1, 2, 3\}$ that do not contain squares of period less than 22.

(854683883 states, 70GB in memory, 2 hours of computation)

A rough estimation seems to indicate that 10^7GB or memory should be enough to remove the second condition on A_i (under the hypothesis that the growth rate is the same)

Other application of this proof technique

A generic sufficient condition for avoidability

Theorem

Let $\mathcal{A}$ be an alphabet and $\mathcal{F}$ be a set of forbiden factors over $\mathcal{A}$. Suppose that there exists a positive real x such that

$$|\mathcal{A}| - \sum_{f \in \mathcal{F}} x^{1-|f|} \geq x \,,$$

then there exists arbitrarily long words over $\mathcal{A}$ that avoid $\mathcal{F}$.

A generic sufficient condition for avoidability

Theorem

Let $\mathcal{A}$ be an alphabet and $\mathcal{F}$ be a set of forbiden factors over $\mathcal{A}$. Suppose that there exists a positive real x such that

$$|\mathcal{A}| - \sum_{f \in \mathcal{F}} x^{1-|f|} \geq x \,,$$

then there exists arbitrarily long words over $\mathcal{A}$ that avoid $\mathcal{F}$.

Fix $\mathcal{A}$ and $\mathcal{F}$.

For any n, let $\mathcal{C}_n$ be the set of words of length n over $\mathcal{A}$ that avoid $\mathcal{F}$.

Lemma

Under the Theorem hypothesis, for all $n \in \mathbb{N}$,

$$|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n| \,.$$

$\implies |\mathcal{C}_n| \geq x^n$

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$,

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}|$$

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f.

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

For all $w \in \mathcal{B}_f$,

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

For all $w \in \mathcal{B}_f$, we have $w = vf$ for some $v \in \mathcal{C}_{n+1-|f|}$.

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

For all $w \in \mathcal{B}_f$, we have $w = vf$ for some $v \in \mathcal{C}_{n+1-|f|}$.

$$|\mathcal{B}_f| \leq |C_{n+1-|f|}|$$

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

For all $w \in \mathcal{B}_f$, we have $w = vf$ for some $v \in \mathcal{C}_{n+1-|f|}$.

$$|\mathcal{B}_f| \leq |C_{n+1-|f|}| \leq \frac{|\mathcal{C}_n|}{x^{|f|-1}} .$$

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

For all $w \in \mathcal{B}_f$, we have $w = vf$ for some $v \in \mathcal{C}_{n+1-|f|}$.

$$|\mathcal{B}_f| \leq |C_{n+1-|f|}| \leq \frac{|\mathcal{C}_n|}{x^{|f|-1}} .$$

Finally,

$$|\mathcal{C}_{n+1}| \geq |\mathcal{A}||C_n| - \sum_{f \in \mathcal{F}} \frac{|\mathcal{C}_n|}{x^{|f|-1}}$$

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$.

For all $w \in \mathcal{B}_f$, we have $w = vf$ for some $v \in \mathcal{C}_{n+1-|f|}$.

$$|\mathcal{B}_f| \leq |C_{n+1-|f|}| \leq \frac{|\mathcal{C}_n|}{x^{|f|-1}} .$$

Finally,

$$|\mathcal{C}_{n+1}| \geq |\mathcal{A}||C_n| - \sum_{f \in \mathcal{F}} \frac{|\mathcal{C}_n|}{x^{|f|-1}} = |C_n| \left(|\mathcal{A}| - \sum_{f \in \mathcal{F}} x^{1-|f|} \right)$$

The proof by induction that $|\mathcal{C}_{n+1}| \geq x \cdot |\mathcal{C}_n|$

By the induction hypothesis, for all $j \in \{0, \ldots, n\}$, $|\mathcal{C}_{n-j}| \leq \frac{|\mathcal{C}_n|}{x^j}$.

Let $\mathcal{B} = \{ua : u \in \mathcal{C}_n, a \in \mathcal{A}\} \setminus \mathcal{C}_{n+1}$, then

$$|C_{n+1}| = |\mathcal{A}| \cdot |\mathcal{C}_n| - |\mathcal{B}| \geq |\mathcal{A}| \cdot |\mathcal{C}_n| - \sum_{f \in \mathcal{F}} |\mathcal{B}_f|$$

For any $f \in \mathcal{B}$, let $\mathcal{B}_f$ be the set of words from $\mathcal{B}$ that ends with an occurence of f. Then $\mathcal{B} = \bigcup\limits_{f \in \mathcal{F}} \mathcal{B}_f$ and $|\mathcal{B}| \leq \sum\limits_{f \in \mathcal{F}} |\mathcal{B}_f|$.

For all $w \in \mathcal{B}_f$, we have $w = vf$ for some $v \in \mathcal{C}_{n+1-|f|}$.

$$|\mathcal{B}_f| \leq |C_{n+1-|f|}| \leq \frac{|\mathcal{C}_n|}{x^{|f|-1}} .$$

Finally,

$$|\mathcal{C}_{n+1}| \geq |\mathcal{A}||C_n| - \sum_{f \in \mathcal{F}} \frac{|\mathcal{C}_n|}{x^{|f|-1}} = |C_n| \left(|\mathcal{A}| - \sum_{f \in \mathcal{F}} x^{1-|f|} \right) \geq x|C_n| \quad \square$$

Back to the result

Theorem (Ochem, 2016 and Rosenfeld, 2022)

Let $\mathcal{A}$ be an alphabet and $\mathcal{F}$ be a set of forbiden factors over $\mathcal{A}$. Suppose that there exists a positive real x such that

$$\mathcal{A} - \sum_{f \in \mathcal{F}} x^{1-|f|} \geq x \,,$$

then the number of words of length n avoiding $\mathcal{F}$ is at least x^n.

Back to the result

Theorem (Ochem, 2016 and Rosenfeld, 2022)

Let $\mathcal{A}$ be an alphabet and $\mathcal{F}$ be a set of forbiden factors over $\mathcal{A}$. Suppose that there exists a positive real x such that

$$\mathcal{A} - \sum_{f \in \mathcal{F}} x^{1-|f|} \geq x \,,$$

then the number of words of length n avoiding $\mathcal{F}$ is at least x^n.

Stronger than Miller (2012) and Pavlov (2021).

Back to the result

Theorem (Ochem, 2016 and Rosenfeld, 2022)

Let $\mathcal{A}$ be an alphabet and $\mathcal{F}$ be a set of forbiden factors over $\mathcal{A}$. Suppose that there exists a positive real x such that

$$\mathcal{A} - \sum_{f \in \mathcal{F}} x^{1-|f|} \geq x \,,$$

then the number of words of length n avoiding $\mathcal{F}$ is at least x^n.

Stronger than Miller (2012) and Pavlov (2021).

Corollary

Let $\mathcal{F}$ be a set of forbidden factors that contain at most one factor of each size in $\{5, 6, 7, \ldots\}$ and no shorter factor. Then the number of words of size n avoiding $\mathcal{F}$ over $\{0, 1\}$ is at least

$$\alpha_1^n \geq 1.755^n \,,$$

where α_1 is the largest root of $x^3 - 2x^2 + x - 1$.

Fractional repetitions

Let $\alpha(k, x)$ be the growth of the language of x-free words over the alphabet $\{1, 2, \ldots, k\}$.

Fractional repetitions

Let $\alpha(k, x)$ be the growth of the language of x-free words over the alphabet $\{1, 2, \ldots, k\}$.

Conjecture (Shur, 2009)

For any fixed integer $n \geq 3$ and arbitrarily large integer k the following holds

$$\alpha\left(k, \frac{n}{n-1}\right) = k + 1 - n - \frac{n-1}{k} + O\left(\frac{1}{k^2}\right)$$

$$\alpha\left(k, \frac{n}{n-1}^+\right) = k + 2 - n - \frac{n-1}{k} + O\left(\frac{1}{k^2}\right)$$

Fractional repetitions

Let $\alpha(k, x)$ be the growth of the language of x-free words over the alphabet $\{1, 2, \ldots, k\}$.

Theorem (Rosenfeld, 2020)

For any fixed integer $n \geq 3$ and arbitrarily large integer k the following holds

$$\alpha\left(k, \frac{n}{n-1}\right) \geq k + 1 - n - \frac{n-1}{k} + O\left(\frac{1}{k^2}\right)$$

$$\alpha\left(k, \frac{n}{n-1}^+\right) \geq k + 2 - n - \frac{n-1}{k} + O\left(\frac{1}{k^2}\right)$$

Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:

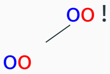
oo

Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:

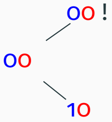


Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:

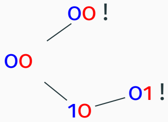


Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:

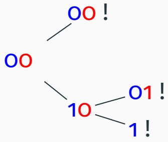


Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:

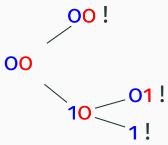


Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:



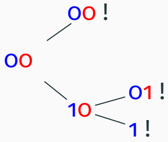
Ben also has a simple strategy with 3 letters.

Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:



Ben also has a simple strategy with 3 letters.

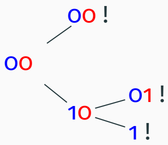
Entropy compression (Grytczuk et al., 2011) $\Longrightarrow$ with ≥ 6 symbols Ann wins.

Thue game

Thue game is between Ann and Ben

- at their turn Ann and Ben add a letter at the end of the word,
- if a square of length at least 4 appears Ben wins.

Over 2 letters:



Ben also has a simple strategy with 3 letters.

Entropy compression (Grytczuk et al., 2011) $\implies$ with ≥ 6 symbols Ann wins.

Theorem (Rosenfeld, 2022)

Over 4 letters Ann has a winning stategy.

The counting argument applied to other objects

A similar technique using formal power series was used by Bell and Goh, Rampersad, Blanchet-Sadri and Woodhouse, Ochem...

The counting argument applied to other objects

A similar technique using formal power series was used by Bell and Goh, Rampersad, Blanchet-Sadri and Woodhouse, Ochem...

This version of the argument is generalizable to tilings, subshifts, hypergraph colorings, graph colorings, k-SAT...

The counting argument applied to other objects

A similar technique using formal power series was used by Bell and Goh, Rampersad, Blanchet-Sadri and Woodhouse, Ochem...

This version of the argument is generalizable to tilings, subshifts, hypergraph colorings, graph colorings, k-SAT...

Theorem (Rosenfeld, 2022)

Let G be a countable amenable group and $\mathcal{F}$ a set of forbidden patterns over the alphabet $\mathcal{A}$. If there exists a positive real β such that,

$$\beta + \sum_{f \in \mathcal{F}} |f| \beta^{1-|f|} \leq |\mathcal{A}|$$

then

$$\alpha(\mathcal{X}_{\mathcal{F}}) \geq \beta\,.$$

Thanks !