Words Avoiding Tangrams

Bartłomiej Pawlik

Silesian University of Technology, Gliwice, Poland

February 25^{th}, 2025

Joint work with M. Dębski, J. Grytczuk, J. Przybyło, and M. Śleszyńska-Nowak

Avoiding squares

Words A, B and C are called *factors* of a word $W = ABC$. Additionally, A is called a *prefix*, and C — a *suffix* of W.

`couscous` `hotshots`

The word W is called *square-free*, if it does not contain any square factor.

`mathematics` `filology`

- Squares are unavoidable in binary words of length greater than 3.
- There exist arbitrarily long square-free ternary words. (Thue, 1906)

Tangrams

Tangram is a word in which each letter appears an even number of times.

Tangrams

Tangram is a word in which each letter appears an even number of times.

The word `tangram` is not a tangram.

Tangrams

Tangram is a word in which each letter appears an even number of times.

The word `tangram` is not a tangram.

Some English tangrams:

`anna`	`appall`	`bilabial`	`bonbon`
`boob`	`bulbul`	`cancan`	`coco`
`couscous`	`dada`	`deed`	`dodo`
`gaga`	`hallah`	`horseshoer`	`mama`
`murmur`	`noon`	`papa`	`peep`
`poop`	`reappear`	`senescence`	`succus`
`tartar`	`tattletale`	`teammate`	`toot`

Is it possible to avoid tangrams in long words?

Let $\mathbb{A} = \{\mathtt{a}_1, \mathtt{a}_2, \ldots, \mathtt{a}_r\}$ be a fixed alphabet.

Every word of length 2^r over $\mathbb{A}$ contains a tangram factor.

Is it possible to avoid tangrams in long words?

Let $\mathbb{A} = \{\mathtt{a}_1, \mathtt{a}_2, \ldots, \mathtt{a}_r\}$ be a fixed alphabet.

Every word of length 2^r over $\mathbb{A}$ contains a tangram factor.

- Let us consider a word $W = \mathtt{w}_1\mathtt{w}_2\cdots\mathtt{w}_{2^r}$.

Is it possible to avoid tangrams in long words?

Let $\mathbb{A} = \{\mathtt{a}_1, \mathtt{a}_2, \ldots, \mathtt{a}_r\}$ be a fixed alphabet.

Every word of length 2^r over $\mathbb{A}$ contains a tangram factor.

- Let us consider a word $W = \mathtt{w}_1\mathtt{w}_2 \cdots \mathtt{w}_{2^r}$.
- For every prefix P_i with length i $(0 \leqslant i \leqslant 2^r)$ of the word W we assign a binary vector $\overrightarrow{v_i} = [v_i(1), v_i(2), \ldots, v_i(r)]$ such that $v_i(j)$ determines the number of occurrences modulo 2 of the letter $\mathtt{a}_j$ w P_i.

Is it possible to avoid tangrams in long words?

Let $\mathbb{A} = \{\mathtt{a}_1, \mathtt{a}_2, \ldots, \mathtt{a}_r\}$ be a fixed alphabet.

Every word of length 2^r over $\mathbb{A}$ contains a tangram factor.

- Let us consider a word $W = \mathtt{w}_1\mathtt{w}_2 \cdots \mathtt{w}_{2^r}$.
- For every prefix P_i with length i $(0 \leqslant i \leqslant 2^r)$ of the word W we assign a binary vector $\overrightarrow{v_i} = [v_i(1), v_i(2), \ldots, v_i(r)]$ such that $v_i(j)$ determines the number of occurrences modulo 2 of the letter $\mathtt{a}_j$ w P_i.
- The number of vectors is $2^r + 1$, the number of different vectors is at most 2^r, so there exists a pair of prefixes P_a, P_b $(0 \leqslant a < b \leqslant 2^r)$ such that $\overrightarrow{v_a} = \overrightarrow{v_b}$

Is it possible to avoid tangrams in long words?

Let $\mathbb{A} = \{\mathtt{a}_1, \mathtt{a}_2, \ldots, \mathtt{a}_r\}$ be a fixed alphabet.

Every word of length 2^r over $\mathbb{A}$ contains a tangram factor.

- Let us consider a word $W = \mathtt{w}_1\mathtt{w}_2\cdots\mathtt{w}_{2^r}$.
- For every prefix P_i with length i $(0 \leqslant i \leqslant 2^r)$ of the word W we assign a binary vector $\overrightarrow{v_i} = [v_i(1), v_i(2), \ldots, v_i(r)]$ such that $v_i(j)$ determines the number of occurrences modulo 2 of the letter $\mathtt{a}_j$ w P_i.
- The number of vectors is $2^r + 1$, the number of different vectors is at most 2^r, so there exists a pair of prefixes P_a, P_b $(0 \leqslant a < b \leqslant 2^r)$ such that $\overrightarrow{v_a} = \overrightarrow{v_b}$, so the factor

$$\mathtt{w}_a\mathtt{w}_{a+1}\cdots\mathtt{w}_{b-1}$$

is a tangram.

Cutting number

Every tangram can be decomposed into factors that can be used to produce two identical words.

Cutting number

Every tangram can be decomposed into factors that can be used to produce two identical words.

If the given tangram is a square, the decomposition is straightforward:

$$\begin{array}{c} \mathtt{coco} \\ \mathtt{co}|\mathtt{co} \\ \swarrow \qquad \searrow \\ \mathtt{co} \qquad\qquad \mathtt{co} \end{array}$$

Cutting number $\mu(T)$ of non-empty tangram T is the least number $k \geqslant 1$ such that

$$T = F_1 F_2 \cdots F_{k+1},$$

where F_i are non-empty words such that

$$F_{\sigma(1)} \cdots F_{\sigma(j)} = F_{\sigma(j+1)} \cdots F_{\sigma(k+1)}$$

for some permutation σ of $\{1, 2, \ldots, k+1\}$ and for some $1 \leqslant j \leqslant k$.

It is convenient to state that $\mu(W) = \infty$ if W is not a tangram.

Let $|W| = n$ be a *length* of the word $W = \mathtt{w}_1\mathtt{w}_2\cdots\mathtt{w}_n$.

Let us notice that for every non-empty tangram T, we have

$$1 \leqslant \mu(T) \leqslant |T| - 1.$$

Non-empty word W is a square if and only if $\mu(W) = 1$.

deed

$$\mu(\texttt{deed}) =$$

$$\texttt{deed}$$

$$\underbrace{\texttt{de}}_{F_1} \mid \underbrace{\texttt{e}}_{F_2} \mid \underbrace{\texttt{d}}_{F_3}$$

$$F_1 = F_3F_2 \qquad \texttt{de} = \texttt{de}$$

$$\mu(\texttt{deed}) =$$

$$\texttt{deed}$$

$$\underbrace{\texttt{de}}_{F_1} \mid \underbrace{\texttt{e}}_{F_2} \mid \underbrace{\texttt{d}}_{F_3}$$

$$F_1 = F_3 F_2 \qquad \texttt{de} = \texttt{de}$$

$$\mu(\texttt{deed}) = 2$$

deed

$$\underbrace{\texttt{de}}_{F_1} \mid \underbrace{\texttt{e}}_{F_2} \mid \underbrace{\texttt{d}}_{F_3}$$

$$F_1 = F_3 F_2 \qquad \texttt{de} = \texttt{de}$$

$$\mu(\texttt{deed}) = 2$$

tattletale

$$\mu(\texttt{tattletale}) =$$

$$\texttt{deed}$$

$$\underbrace{\texttt{de}}_{F_1} \mid \underbrace{\texttt{e}}_{F_2} \mid \underbrace{\texttt{d}}_{F_3}$$

$$F_1 = F_3F_2 \qquad \texttt{de} = \texttt{de}$$

$$\mu(\texttt{deed}) = 2$$

$$\texttt{tattletale}$$

$$\underbrace{\texttt{tat}}_{F_1} \mid \underbrace{\texttt{tle}}_{F_2} \mid \underbrace{\texttt{ta}}_{F_3} \mid \underbrace{\texttt{le}}_{F_4}$$

$$F_1F_4 = F_3F_2 \qquad \texttt{tatle} = \texttt{tatle}$$

$$\mu(\texttt{tattletale}) = 3$$

$\mu(T) = 1$ for:
`bonbon`, `bulbul`, `cancan`, `coco`, `couscous`, `dada`, `dodo`, `gaga`, `mama`, `murmur`, `papa`, `tartar`

$\mu(T) = 2$ for:
`anna`, `boob`, `deed`, `noon`, `peep`, `poop`, `toot`

$\mu(T) = 3$ for
`appall`, `bilabial`, `hallah`, `reappear`, `succus`, `tattletale`, `teammate`

$\mu(T) = 4$ for
`horseshoer`, `senescence`

The word W is k-tangram-free, if it does not contain a factor F such that $\mu(F) \leqslant k$.

The word W is k-tangram-free, if it does not contain a factor F such that $\mu(F) \leqslant k$.

For example, the word W is 1-tangram-free if and only if it is square-free.

The word W is k-tangram-free, if it does not contain a factor F such that $\mu(F) \leqslant k$.

For example, the word W is 1-tangram-free if and only if it is square-free.

The smallest size of the alphabet for which there exists an arbitrarily long k-tangram-free word is called the *tangram-free size* and is denoted by $t(k)$.

The word W is k-tangram-free, if it does not contain a factor F such that $\mu(F) \leqslant k$.

For example, the word W is 1-tangram-free if and only if it is square-free.

The smallest size of the alphabet for which there exists an arbitrarily long k-tangram-free word is called the *tangram-free size* and is denoted by $t(k)$.

- $t(1) = 3$

The word W is k-tangram-free, if it does not contain a factor F such that $\mu(F) \leqslant k$.

For example, the word W is 1-tangram-free if and only if it is square-free.

The smallest size of the alphabet for which there exists an arbitrarily long k-tangram-free word is called the *tangram-free size* and is denoted by $t(k)$.

- $t(1) = 3$
- $t(2) = 3$

$t(2) = 3$

Let W be the infinite ternary square-free word.

$t(2) = 3$

Let W be the infinite ternary square-free word.
Let us assume that T is a tangram factor of W such that $\mu(T) = 2$

$t(2) = 3$

Let W be the infinite ternary square-free word.
Let us assume that T is a tangram factor of W such that $\mu(T) = 2$, so

$$T = F_1 F_2 F_3$$

for some non-empty factors F_1, F_2, F_3 that satisfy a certain equation.

$t(2) = 3$

Let W be the infinite ternary square-free word.
Let us assume that T is a tangram factor of W such that $\mu(T) = 2$, so

$$T = F_1 F_2 F_3$$

for some non-empty factors F_1, F_2, F_3 that satisfy a certain equation.

We show that for every possible equation between segments, a square appears in T, which gives us a contradiction with the fact that there are no squares in the word W.

$t(2) = 3$

Let W be the infinite ternary square-free word.
Let us assume that T is a tangram factor of W such that $\mu(T) = 2$, so

$$T = F_1F_2F_3$$

for some non-empty factors F_1, F_2, F_3 that satisfy a certain equation.

We show that for every possible equation between segments, a square appears in T, which gives us a contradiction with the fact that there are no squares in the word W.

eq.	T	XX
$F_1 = F_2F_3$	$F_2F_3F_2F_3$	$F_2F_3F_2F_3$
$F_1 = F_3F_2$	$F_3F_2F_2F_3$	F_2F_2
$F_2 = F_1F_3$	$F_1F_1F_3F_3$	$F_1F_1,\ F_3F_3$
$F_2 = F_3F_1$	$F_1F_3F_1F_3$	$F_1F_3F_1F_3$
$F_3 = F_1F_2$	$F_1F_2F_1F_2$	$F_1F_2F_1F_2$
$F_3 = F_2F_1$	$F_1F_2F_2F_1$	F_2F_2

- $t(1) = t(2) = 3$ (Thue's theorem)
- $t(3) = t(4) = 4$ (P. Ochem, T. Pierron, 2025+)
- $t(k) = ?$

- $t(1) = t(2) = 3$ (Thue's theorem)
- $t(3) = t(4) = 4$ (P. Ochem, T. Pierron, 2025+)
- $t(k) =?$

Theorem 2 (Dębski, Grytczuk, Przybyło, P., Śleszyńska-Nowak, 2024)

For every $k \geqslant 4$ we have

$$t(k) \leqslant k + 1.$$

$t(k) \leqslant k+1$

In the word

$$\texttt{reappear},$$

the *distance* between the factors `app` and `pear` is 2.

$t(k) \leqslant k+1$

In the word

`reappear`,

the *distance* between the factors `app` and `pear` is 2.

Theorem 1 is a consequence of the following facts:

Theorem (Dejean conjecture, 1972-2009)

For every $r \geqslant 5$ there is an arbitrarily long word D (the Dejean word) over the alphabet of size r such that the distance between two consecutive occurences of F in D is at least $(r-1)|F|$.

$t(k) \leqslant k + 1$

In the word

`reappear`,

the *distance* between the factors `app` and `pear` is 2.

Theorem 1 is a consequence of the following facts:

Theorem (Dejean conjecture, 1972-2009)

For every $r \geqslant 5$ there is an arbitrarily long word D (the Dejean word) over the alphabet of size r such that the distance between two consecutive occurences of F in D is at least $(r-1)|F|$.

Theorem (Dębski, Grytczuk, P., Przybyło, Śleszyńska-Nowak, 2024)

Every factor F in the Dejean word D over the alphabet of the size $r \geqslant 5$ satisfies $\mu(F) \geqslant r$.

$$t(k) = \mathcal{O}(\log_2 k)$$

For every $k > 20\,478$, we have a better estimate for $t(k)$:

$t(k) = \mathcal{O}(\log_2 k)$

For every $k > 20\,478$, we have a better estimate for $t(k)$:

Theorem 1 (Dębski, Grytczuk, P., Przybyło, Śleszyńska-Nowak, 2024)

For every $k \geqslant 3$, we have $t(k) \leqslant 1024 \cdot \lceil \log_2 k + \log_2 \log_2 k + 1 \rceil$.

$$t(k) = \Theta(\log_2 k)$$

Proposition (Dębski, Grytczuk, P., Przybyło, Śleszyńska-Nowak, 2024)

For every $k \geqslant 1$, we have $t(k) \geqslant \log_2(k+2)$.

$t(k) = \Theta(\log_2 k)$

Proposition (Dębski, Grytczuk, P., Przybyło, Śleszyńska-Nowak, 2024)

For every $k \geqslant 1$, we have $t(k) \geqslant \log_2(k+2)$.

- Suppose that $t(k) < \log_2(k+2)$.

$t(k) = \Theta(\log_2 k)$

Proposition (Dębski, Grytczuk, P., Przybyło, Śleszyńska-Nowak, 2024)

For every $k \geqslant 1$, we have $t(k) \geqslant \log_2(k+2)$.

- Suppose that $t(k) < \log_2(k+2)$.
- Thus, there exists an arbitrarily long word W over an alphabet of size $t(k) = q < \log_2(k+2)$, such that all factors F of W satisfy $\mu(F) \geqslant k+1$.

$t(k) = \Theta(\log_2 k)$

Proposition (Dębski, Grytczuk, P., Przybyło, Śleszyńska-Nowak, 2024)

For every $k \geqslant 1$, we have $t(k) \geqslant \log_2(k+2)$.

- Suppose that $t(k) < \log_2(k+2)$.
- Thus, there exists an arbitrarily long word W over an alphabet of size $t(k) = q < \log_2(k+2)$, such that all factors F of W satisfy $\mu(F) \geqslant k+1$.
- Let $|W| = 2^q$. Then W contains the tangram factor T.

$t(k) = \Theta(\log_2 k)$

Proposition (Dębski, Grytczuk, P., Przybyło, Śleszyńska-Nowak, 2024)

For every $k \geqslant 1$, we have $t(k) \geqslant \log_2(k+2)$.

- Suppose that $t(k) < \log_2(k+2)$.
- Thus, there exists an arbitrarily long word W over an alphabet of size $t(k) = q < \log_2(k+2)$, such that all factors F of W satisfy $\mu(F) \geqslant k+1$.
- Let $|W| = 2^q$. Then W contains the tangram factor T.
- We have the following contradiction with $\mu(T) \geqslant k+1$:

$$\mu(T) \leqslant |T| - 1 \leqslant |W| - 1 = 2^q - 1 < 2^{\log_2(k+2)} - 1 = k+1.$$

Question: Constant C

We already know that $t(k) = \Theta(\log_2 k)$. How close is actually $t(k)$ to the function $\log_2 k$?

Question: Constant C

We already know that $t(k) = \Theta(\log_2 k)$. How close is actually $t(k)$ to the function $\log_2 k$?

Problem 1

Is there a constant C such that $t(k) \leqslant \log_2 k + C$?

Abelian case

The word W is an *abelian square* if $W = XX'$, where X' is an anagram of X.

Abelian case

The word W is an *abelian square* if $W = XX'$, where X' is an anagram of X.

Every square is an abelian square.

`intestines` `teammate` `signings`

Abelian case

The word W is an *abelian square* if $W = XX'$, where X' is an anagram of X.

Every square is an abelian square.

`intestines` `teammate` `signings`

What is the minimum size of an alphabet such that there exists an arbitrarily long word without abelian squares as factors?

Abelian case

The word W is an *abelian square* if $W = XX'$, where X' is an anagram of X.

Every square is an abelian square.

`intestines` `teammate` `signings`

What is the minimum size of an alphabet such that there exists an arbitrarily long word without abelian squares as factors?

Theorem (V. Keränen, 2005)

Abelian squares are avoidable on 4 letters.

Goldberg-West Theorem - binary case (1985)

For every binary tangram B, there exist words X, V, Y such that $B = XVY$ and the words V and XY have the same number of occurrences of the symbols 0 and 1.

Goldberg-West Theorem - binary case (1985)

For every binary tangram B, there exist words X, V, Y such that $B = XVY$ and the words V and XY have the same number of occurrences of the symbols 0 and 1.

0 0 1 0 1 1 0 0 0 0 1 0 1 1 1 0 0 1

Goldberg-West Theorem - binary case (1985)

For every binary tangram B, there exist words X, V, Y such that $B = XVY$ and the words V and XY have the same number of occurrences of the symbols 0 and 1.

0 0 1 0 1 1 0 0 0 0 1 0 1 1 1 0 0 1

	V	XY
number of 0's		
number of 1's		

Goldberg-West Theorem - binary case (1985)

For every binary tangram B, there exist words X, V, Y such that $B = XVY$ and the words V and XY have the same number of occurrences of the symbols 0 and 1.

|0 0 1 0 1 1 0 0 0|0 1 0 1 1 1 0 0 1

	V	XY
number of 0's	6	4
number of 1's	3	5

Goldberg-West Theorem - binary case (1985)

For every binary tangram B, there exist words X, V, Y such that $B = XVY$ and the words V and XY have the same number of occurrences of the symbols 0 and 1.

0|0 1 0 1 1 0 0 0 0|1 0 1 1 1 0 0 1

	V	XY
number of 0's	6	4
number of 1's	3	5

Goldberg-West Theorem - binary case (1985)

For every binary tangram B, there exist words X, V, Y such that $B = XVY$ and the words V and XY have the same number of occurrences of the symbols 0 and 1.

0 0|1 0 1 1 0 0 0 0 1|0 1 1 1 0 0 1

	V	XY
number of 0's	5	5
number of 1's	4	4

The necklace splitting theorem

Suppose a necklace has $k \cdot n$ beads, chosen form t different colors, and that there are $k \cdot a_i$ beads of color i, $1 \leqslant i \leqslant t$.

A *k-splitting* if the necklace is a partition of the necklace into k parts, each consisting of a finite number of nonoverlapping intervals of beads, whose union captures precisely a_i beads of color i, $1 \leqslant i \leqslant t$.

The *size* of the k-splitting is the number of cuts that form the intervals of the splitting.

The necklace splitting theorem

Every necklace with $k \cdot a_i$ beads of color i, $1 \leqslant i \leqslant t$, has a k-splitting of size at most $(k-1) \cdot t$. The number $(k-1) \cdot t$ is the best possible.

Avoiding anagrams

Let $\alpha(T)$ be the least number of cuts needed to decompose given tangram T into factors that can be made into a pair of anagrams.

Avoiding anagrams

Let $\alpha(T)$ be the least number of cuts needed to decompose given tangram T into factors that can be made into a pair of anagrams.

$$\alpha(\texttt{abcacb}) = 1 \qquad \alpha(\texttt{aabbcc}) = 3$$

Avoiding anagrams

Let $\alpha(T)$ be the least number of cuts needed to decompose given tangram T into factors that can be made into a pair of anagrams.

$$\alpha(\mathtt{abcacb}) = 1 \qquad \alpha(\mathtt{aabbcc}) = 3$$

We have $\alpha(T) \leqslant q$.
(q - size of an alphabet, splitting necklace theorem)

Avoiding anagrams

Let $\alpha(T)$ be the least number of cuts needed to decompose given tangram T into factors that can be made into a pair of anagrams.

$$\alpha(\texttt{abcacb}) = 1 \qquad \alpha(\texttt{aabbcc}) = 3$$

We have $\alpha(T) \leqslant q$.
(q - size of an alphabet, splitting necklace theorem)

Let k-anagram-free word be the word W such that every tangram factor F of W satisfy $\alpha(F) \geqslant k + 1$.

Let $a(k)$ denote the least size of an alphabet needed to construct arbitrarily long k-anagram-free word.

Avoiding anagrams

Let $\alpha(T)$ be the least number of cuts needed to decompose given tangram T into factors that can be made into a pair of anagrams.

$$\alpha(\texttt{abcacb}) = 1 \qquad \alpha(\texttt{aabbcc}) = 3$$

We have $\alpha(T) \leqslant q$.
(q - size of an alphabet, splitting necklace theorem)

Let k-anagram-free word be the word W such that every tangram factor F of W satisfy $\alpha(F) \geqslant k+1$.

Let $a(k)$ denote the least size of an alphabet needed to construct arbitrarily long k-anagram-free word.

$a(1) = 4$ (Theorem of Keränen).

Avoiding anagrams

Let $\alpha(T)$ be the least number of cuts needed to decompose given tangram T into factors that can be made into a pair of anagrams.

$$\alpha(\texttt{abcacb}) = 1 \qquad \alpha(\texttt{aabbcc}) = 3$$

We have $\alpha(T) \leqslant q$.
(q - size of an alphabet, splitting necklace theorem)

Let k-anagram-free word be the word W such that every tangram factor F of W satisfy $\alpha(F) \geqslant k+1$.

Let $a(k)$ denote the least size of an alphabet needed to construct arbitrarily long k-anagram-free word.

$a(1) = 4$ (Theorem of Keränen).

Conjecture (N. Alon, J. Grytczuk, M. Michałek, M. Lasoń, 2009)

For every $k \geqslant 1$, we have $a(k) \leqslant k+3$.

Cutting shuffle squares

A *shuffle square* is a finite word that can be formed by self-shuffling a word.

Cutting shuffle squares

A *shuffle square* is a finite word that can be formed by self-shuffling a word.
For instance, the word

acaece

is a shuffle square:

acaece → ace, ace.

Cutting shuffle squares

A *shuffle square* is a finite word that can be formed by self-shuffling a word.
For instance, the word

$$\mathtt{acaece}$$

is a shuffle square:

$$\mathtt{acaece} \;\rightarrow\; \mathtt{ace},\ \mathtt{ace}.$$

Let $n \geqslant 1$. The word $W = \mathtt{w}_1\mathtt{w}_2 \cdots \mathtt{w}_{2n}$ is *a shuffle square*, if there exist index sets

$$I = \{i_1, i_2, \ldots, i_n\}, \quad J = \{j_1, j_2, \ldots, j_n\}$$

such that

$$i_1 < i_2 < \ldots < i_n, \quad j_1 < j_2 < \ldots < j_n,$$

$I \cap J = \emptyset$ and $\mathtt{w}_{i_r} = \mathtt{w}_{j_r}$ for $1 \leqslant r \leqslant n$.

Not every tangram is a shuffle square!

001100 011110

Not every tangram is a shuffle square!

001100 011110

001100 011110

Not every tangram is a shuffle square!

001100 011110

001100 011110

Binary tangrams with length 6:

000000	010010	111111	101101
000011	010100	111100	101011
000101	011000	111010	100111
000110	100001	111001	011110
001001	100010	110110	011101
001010	100100	110101	011011
001100	101000	110011	010111
010001	110000	101110	001111

Avoiding shuffle squares

Size of an alphabet on which we can avoid shuffle-squares in arbitrarily long words:

Avoiding shuffle squares

Size of an alphabet on which we can avoid shuffle-squares in arbitrarily long words:

10^{40}	2014	J. Currie

Avoiding shuffle squares

Size of an alphabet on which we can avoid shuffle-squares in arbitrarily long words:

10^{40}	2014	J. Currie
10	2015	M. Müller

Avoiding shuffle squares

Size of an alphabet on which we can avoid shuffle-squares in arbitrarily long words:

10^{40}	2014	J. Currie
10	2015	M. Müller
7	2016	G. Guégan, P. Ochem
7	2016	J. Grytczuk, J. Kozik, B. Zaleski

Avoiding shuffle squares

Size of an alphabet on which we can avoid shuffle-squares in arbitrarily long words:

10^{40}	2014	J. Currie
10	2015	M. Müller
7	2016	G. Guégan, P. Ochem
7	2016	J. Grytczuk, J. Kozik, B. Zaleski
6	2023	L. Bulteau, V. Jugé, S. Vialette

Avoiding shuffle squares

Size of an alphabet on which we can avoid shuffle-squares in arbitrarily long words:

10^{40}	2014	J. Currie
10	2015	M. Müller
7	2016	G. Guégan, P. Ochem
7	2016	J. Grytczuk, J. Kozik, B. Zaleski
6	2023	L. Bulteau, V. Jugé, S. Vialette

4? 5? 6?

The words V and W are *conjugate* if there exist words A and B such that $V = AB$ and $W = BA$.

The words V and W are *conjugate* if there exist words A and B such that $V = AB$ and $W = BA$.

$$
\begin{array}{lcllcl}
\underline{0}00110 & \longrightarrow & 001100 & \qquad \underline{111}001 & \longrightarrow & 001111 \\
\underline{01}0001 & \longrightarrow & 000101 & \qquad \underline{10}1110 & \longrightarrow & 111010 \\
\underline{0}11000 & \longrightarrow & 110000 & \qquad \underline{1}00111 & \longrightarrow & 001111 \\
\underline{1}00001 & \longrightarrow & 000011 & \qquad \underline{0}11110 & \longrightarrow & 111100 \\
\underline{1}00010 & \longrightarrow & 000101 & \qquad \underline{0}11101 & \longrightarrow & 111010
\end{array}
$$

Conjecture 1 (Grytczuk, P., Pleszczyński, 2023)

For every binary tangram W, there exists a shuffle square S such that W and S are conjugate.

Conjecture 1 (Grytczuk, P., Pleszczyński, 2023)

For every binary tangram W, there exists a shuffle square S such that W and S are conjugate.

Results:

Proposition (Grytczuk, P., Pleszczyński, 2023)

For every binary tangram W with four 1's there exists a shuffle square S such that W and S are conjugate.

Theorem (Grytczuk, P., Pleszczyński, 2023)

Every binary tangram can be decomposed into words A and B such that A and B are conjugate.

Conjecture 1 (Grytczuk, P., Pleszczyński, 2023)

For every binary tangram W, there exists a shuffle square S such that W and S are conjugate.

Results:

Proposition (Grytczuk, P., Pleszczyński, 2023)

For every binary tangram W with four 1's there exists a shuffle square S such that W and S are conjugate.

Theorem (Grytczuk, P., Pleszczyński, 2023)

Every binary tangram can be decomposed into words A and B such that A and B are conjugate.

Counterexample:

$$000001001111000011101111$$

(False) Conjecture 1 may be stated as follows:

Conjecture 1

Every binary tangram can be *cut* into two words A and B such that at least one of two rearrangements of A and B (AB or BA) forms a shuffle square.

(False) Conjecture 1 may be stated as follows:

Conjecture 1

Every binary tangram can be *cut* into two words A and B such that at least one of two rearrangements of A and B (AB or BA) forms a shuffle square.

$$011000 \; \rightarrow 011|000 \; \rightarrow 000|011$$

We state the following

Conjecture 2

Every binary tangram can be *cut* with two cuts into three words A, B, C such that at least one of the six possible rearrangements of A, B, and C forms a shuffle square.

We state the following

Conjecture 2

Every binary tangram can be *cut* with two cuts into three words A, B, C such that at least one of the six possible rearrangements of A, B, and C forms a shuffle square.

000001001111000011101111

$\downarrow$

0000|01|001111000011101111

$\downarrow$

0000|001111000011101111|01

We state the following

Conjecture 2

Every binary tangram can be *cut* with two cuts into three words A, B, C such that at least one of the six possible rearrangements of A, B, and C forms a shuffle square.

0000010011110000011101111

$\downarrow$

0000|01|00111100000111101111

$\downarrow$

0000|00111100000111101111|01

Conjecture 3

Every k-ary tangram can be *cut* with k cuts into $(k+1)$ words $A_1, A_2, \ldots, A_{k+1}$ such that at least one of the rearrangements of $A_1, \ldots, A_{k+1}$ forms a shuffle square.

Thank you!