On the number of squares in a word

Shuo LI
Joint work with Srečko Brlek, Francesco Dolce, Élise Vandomme

Université du Québec à Montréal

September 19 2022

Presentation plan

1. Definitions and background
2. Proof of the new result
3. Discussion
4. Square defect and some conjectures

Definitions

- Let $w = w[1]w[2]...w[n]$ be a word or a string.
  n is called the length of w. It is denoted by $|w|$.
- The alphabet of w is the collection of distinct letters in w.
  It is denoted by $\mathrm{Alph}(w)$.
- The factors of w are words of the type $w[i]w[i+1]...w[j]$, with $1 \leq i \leq j \leq |w|$. Let $\mathrm{Fact}(w)$ denote the set of factors of w.

Definitions

- Let $w = w[1]w[2]...w[n]$ be a word or a string.
  n is called the length of w. It is denoted by $|w|$.
- The alphabet of w is the collection of distinct letters in w.
  It is denoted by $\mathrm{Alph}(w)$.
- The factors of w are words of the type $w[i]w[i+1]...w[j]$, with $1 \le i \le j \le |w|$. Let $\mathrm{Fact}(w)$ denote the set of factors of w.

Definitions

- Let $w = w[1]w[2]...w[n]$ be a word or a string.
  n is called the length of w. It is denoted by $|w|$.
- The alphabet of w is the collection of distinct letters in w. It is denoted by $\text{Alph}(w)$.
- The factors of w are words of the type $w[i]w[i+1]...w[j]$, with $1 \leq i \leq j \leq |w|$. Let $\text{Fact}(w)$ denote the set of factors of w.

Patterns

- A palindrome is a word $w = w[1]w[2]...w[n]$ such that

$$w[1]w[2]...w[n] = w[n]w[n-1]...w[1].$$

- A σ-palindrome is a word $w = w[1]w[2]...w[n]$ such that

$$w[1]w[2]...w[n] = \sigma(w[n]w[n-1]...w[1]),$$

where σ is an involution on $\mathrm{Alph}(w)$ satisfying $\sigma^2 = Id$.

- A square is a word $w = w[1]w[2]...w[2n]$ such that

$$w[1]w[2]...w[n] = w[n+1]w[n+2]...w[2n].$$

Patterns

- A palindrome is a word $w = w[1]w[2]...w[n]$ such that

$$w[1]w[2]...w[n] = w[n]w[n-1]...w[1].$$

- A σ-palindrome is a word $w = w[1]w[2]...w[n]$ such that

$$w[1]w[2]...w[n] = \sigma(w[n]w[n-1]...w[1]),$$

where σ is an involution on $\text{Alph}(w)$ satisfying $\sigma^2 = Id$.

- A square is a word $w = w[1]w[2]...w[2n]$ such that

$$w[1]w[2]...w[n] = w[n+1]w[n+2]...w[2n].$$

Patterns

- A palindrome is a word $w = w[1]w[2]...w[n]$ such that

$$w[1]w[2]...w[n] = w[n]w[n-1]...w[1].$$

- A σ-palindrome is a word $w = w[1]w[2]...w[n]$ such that

$$w[1]w[2]...w[n] = \sigma(w[n]w[n-1]...w[1]),$$

where σ is an involution on $\mathrm{Alph}(w)$ satisfying $\sigma^2 = Id$.

- A square is a word $w = w[1]w[2]...w[2n]$ such that

$$w[1]w[2]...w[n] = w[n+1]w[n+2]...w[2n].$$

Counting patterns in a word

For a finite word w, let $p(w)$, $p_\sigma(w)$ and $s(w)$ denote respectively the number of distinct palindromes, σ-palindromes and squares (counting the empty word) in w.

- Palindrome : in a given position i, there exists one new palindromic suffix.(Droubay, Justin and Pirillo, 2001)

$$p(w) \leq |w| + 1.$$

- σ-palindrome : in a given position i, there exists one new σ-palindromic suffix.(Blondin-Masse, Brlek, Garon, and Labbe 2008, Pelantová and Starosta, 2014)

$$p_\sigma(w) \leq |w| + 1.$$

- Square : in a given position i, there can be two new square suffixes. (Fraenkel and Simpson, 1998)

$$s(w) \leq 2|w| + 1.$$

Counting patterns in a word

For a finite word w, let $p(w)$, $p_\sigma(w)$ and $s(w)$ denote respectively the number of distinct palindromes, σ-palindromes and squares (counting the empty word) in w.

- Palindrome : in a given position i, there exists one new palindromic suffix.(Droubay, Justin and Pirillo, 2001)

$$p(w) \leq |w| + 1.$$

- σ-palindrome : in a given position i, there exists one new σ-palindromic suffix.(Blondin-Masse, Brlek, Garon, and Labbe 2008, Pelantová and Starosta, 2014)

$$p_\sigma(w) \leq |w| + 1.$$

- Square : in a given position i, there can be two new square suffixes. (Fraenkel and Simpson, 1998)

$$s(w) \leq 2|w| + 1.$$

Counting patterns in a word

For a finite word w, let $p(w)$, $p_\sigma(w)$ and $s(w)$ denote respectively the number of distinct palindromes, σ-palindromes and squares (counting the empty word) in w.

- Palindrome : in a given position i, there exists one new palindromic suffix.(Droubay, Justin and Pirillo, 2001)

$$p(w) \leq |w| + 1.$$

- σ-palindrome : in a given position i, there exists one new σ-palindromic suffix.(Blondin-Masse, Brlek, Garon, and Labbe 2008, Pelantová and Starosta, 2014)

$$p_\sigma(w) \leq |w| + 1.$$

- Square : in a given position i, there can be two new square suffixes. (Fraenkel and Simpson, 1998)

$$s(w) \leq 2|w| + 1.$$

An example of words with many squares

- First consideration from Fraenkel and Simpson(1998).
- For $i \geq 1$, let $q_i = 0^{i+1}10^i10^{i+1}1$ and let $w_i = q_1q_2...q_i$.
- Let $s^*(w)$ denote the number of distinct nonempty squares in w.

Theorem (Fraenkel,Simpson(1998))

For the sequence $(w_i)_{i\in\mathbb{N}}$ defined as above $|w_i| = (3i^2 + 13i)/2$ and $s^*(w_i) = 3i^2/2 + \lfloor (i+1)/2 \rfloor + 7i/2 - 3$.

- Observation :

$$s^*(w_i) = |w_i| - o(\sqrt{|w_i|}).$$

An example of words with many squares

- First consideration from Fraenkel and Simpson(1998).
- For $i \geq 1$, let $q_i = 0^{i+1}10^i10^{i+1}1$ and let $w_i = q_1q_2...q_i$.
- Let $s^*(w)$ denote the number of distinct nonempty squares in w.

Theorem (Fraenkel,Simpson(1998))

For the sequence $(w_i)_{i\in\mathbb{N}}$ defined as above $|w_i| = (3i^2+13i)/2$ and $s^*(w_i) = 3i^2/2 + \lfloor (i+1)/2 \rfloor + 7i/2 - 3$.

- Observation :

$$s^*(w_i) = |w_i| - o(\sqrt{|w_i|}).$$

An example of words with many squares

- First consideration from Fraenkel and Simpson(1998).
- For $i \geq 1$, let $q_i = 0^{i+1}10^i10^{i+1}1$ and let $w_i = q_1 q_2 ... q_i$.
- Let $s^*(w)$ denote the number of distinct nonempty squares in w.

Theorem (Fraenkel,Simpson(1998))

For the sequence $(w_i)_{i \in \mathbf{N}}$ defined as above $|w_i| = (3i^2 + 13i)/2$ and $s^*(w_i) = 3i^2/2 + \lfloor (i+1)/2 \rfloor + 7i/2 - 3$.

- Observation :

$$s^*(w_i) = |w_i| - o(\sqrt{|w_i|}).$$

An example of words with many squares

- First consideration from Fraenkel and Simpson(1998).
- For $i \geq 1$, let $q_i = 0^{i+1}10^i10^{i+1}1$ and let $w_i = q_1 q_2 \ldots q_i$.
- Let $s^*(w)$ denote the number of distinct nonempty squares in w.

Theorem (Fraenkel,Simpson(1998))

For the sequence $(w_i)_{i \in \mathbf{N}}$ *defined as above* $|w_i| = (3i^2 + 13i)/2$ *and* $s^*(w_i) = 3i^2/2 + \lfloor (i+1)/2 \rfloor + 7i/2 - 3$.

- Observation :

$$s^*(w_i) = |w_i| - o(\sqrt{|w_i|}).$$

An example of words with many squares

- First consideration from Fraenkel and Simpson(1998).
- For $i \geq 1$, let $q_i = 0^{i+1}10^i10^{i+1}1$ and let $w_i = q_1q_2...q_i$.
- Let $s^*(w)$ denote the number of distinct nonempty squares in w.

Theorem (Fraenkel,Simpson(1998))

For the sequence $(w_i)_{i\in\mathbf{N}}$ defined as above $|w_i| = (3i^2 + 13i)/2$ and $s^*(w_i) = 3i^2/2 + \lfloor (i+1)/2 \rfloor + 7i/2 - 3$.

- Observation :

$$s^*(w_i) = |w_i| - o(\sqrt{|w_i|}).$$

Fraenkel-Simpson's conjecture

Conjecture (Fraenkel, Simpson (1998))

For any finite word w, one has

$$s^*(w) \leq |w|.$$

- Results : Any finite word w contains at least at most :
- Fraenkel, Simpson (1998) : $s^*(w) \leq 2|w|$;
- Ilie (2005) : $s^*(w) \leq 2|w| - o(log(|w|))$;
- Lam (2013) : $s^*(w) \leq 95/48|w|$;
- Deza, Franek and Thierry (2015) : $s^*(w) \leq 11/6|w|$;
- Thierry (2020) : $s^*(w) \leq 3/2|w|$.

Fraenkel-Simpson's conjecture

Conjecture (Fraenkel, Simpson (1998))

For any finite word w, one has

$$s^*(w) \leq |w|.$$

- Results : Any finite word w contains at least at most :
  - Fraenkel, Simpson (1998) : $s^*(w) \leq 2|w|$;
  - Ilie (2005) : $s^*(w) \leq 2|w| - o(log(|w|))$;
  - Lam (2013) : $s^*(w) \leq 95/48|w|$;
  - Deza, Franek and Thierry (2015) : $s^*(w) \leq 11/6|w|$;
  - Thierry (2020) : $s^*(w) \leq 3/2|w|$.

Fraenkel-Simpson's conjecture

Conjecture (Fraenkel, Simpson (1998))

For any finite word w, one has

$$s^*(w) \leq |w|.$$

- Results : Any finite word w contains at least at most :
- Fraenkel, Simpson (1998) : $s^*(w) \leq 2|w|$;
- Ilie (2005) : $s^*(w) \leq 2|w| - o(log(|w|))$;
- Lam (2013) : $s^*(w) \leq 95/48|w|$;
- Deza, Franek and Thierry (2015) : $s^*(w) \leq 11/6|w|$;
- Thierry (2020) : $s^*(w) \leq 3/2|w|$.

Fraenkel-Simpson's conjecture

Conjecture (Fraenkel, Simpson (1998))

For any finite word w, one has

$$s^*(w) \leq |w|.$$

- Results : Any finite word w contains at least at most :
- Fraenkel, Simpson (1998) : $s^*(w) \leq 2|w|$;
- Ilie (2005) : $s^*(w) \leq 2|w| - o(log(|w|))$;
- Lam (2013) : $s^*(w) \leq 95/48|w|$;
- Deza, Franek and Thierry (2015) : $s^*(w) \leq 11/6|w|$;
- Thierry (2020) : $s^*(w) \leq 3/2|w|$.

Fraenkel-Simpson's conjecture

Conjecture (Fraenkel, Simpson (1998))

For any finite word w, one has

$$s^*(w) \leq |w|.$$

- Results : Any finite word w contains at least at most :
- Fraenkel, Simpson (1998) : $s^*(w) \leq 2|w|$;
- Ilie (2005) : $s^*(w) \leq 2|w| - o(log(|w|))$;
- Lam (2013) : $s^*(w) \leq 95/48|w|$;
- Deza, Franek and Thierry (2015) : $s^*(w) \leq 11/6|w|$;
- Thierry (2020) : $s^*(w) \leq 3/2|w|$.

Fraenkel-Simpson's conjecture

Conjecture (Fraenkel, Simpson (1998))

For any finite word w, one has

$$s^*(w) \leq |w|.$$

- Results : Any finite word w contains at least at most :
- Fraenkel, Simpson (1998) : $s^*(w) \leq 2|w|$;
- Ilie (2005) : $s^*(w) \leq 2|w| - o(log(|w|))$;
- Lam (2013) : $s^*(w) \leq 95/48|w|$;
- Deza, Franek and Thierry (2015) : $s^*(w) \leq 11/6|w|$;
- Thierry (2020) : $s^*(w) \leq 3/2|w|$.

Fraenkel-Simpson's conjecture

Conjecture (Fraenkel, Simpson (1998))

For any finite word w, one has

$$s^*(w) \leq |w|.$$

- Results : Any finite word w contains at least at most :
- Fraenkel, Simpson (1998) : $s^*(w) \leq 2|w|$;
- Ilie (2005) : $s^*(w) \leq 2|w| - o(log(|w|))$;
- Lam (2013) : $s^*(w) \leq 95/48|w|$;
- Deza, Franek and Thierry (2015) : $s^*(w) \leq 11/6|w|$;
- Thierry (2020) : $s^*(w) \leq 3/2|w|$.

New result

Theorem (Brlek, Li (2022))

Any finite word $w = w[1]w[2]...w[n]$, if we let $|\mathrm{Alph}(w)|$ denote the number of distinct letters in w, one has

$$s^*(w) \leq |w| - |\mathrm{Alph}(w)|.$$

Moreover, if we count the empty word, then one has

$$s(w) \leq |w| - |\mathrm{Alph}(w)| + 1.$$

Rauzy graphs

- Definitions : Let w be a finite word. For any integer i satisfying $1 \leq i \leq |w|$, let $L_w(i)$ denote the set of length-i factors of w and let $C_w(i)$ be the cardinality of $L_w(i)$.
- The i-th Rauzy graph of w is $\Gamma_w(i) = \{v_w(i), e_w(i)\}$ such that
  (i) $v_w(i) = L_w(i); e_w(i) = L_w(i+1)$,
  (ii)$e \in e_w(i+1)$ from v_1 to v_2 iff there exist $\alpha, \beta \in \mathrm{Alph}(w), y \in L_w(i-1)$ such that $v_1 = \alpha y, v_2 = y\beta, e = \alpha y \beta$.

Rauzy graphs

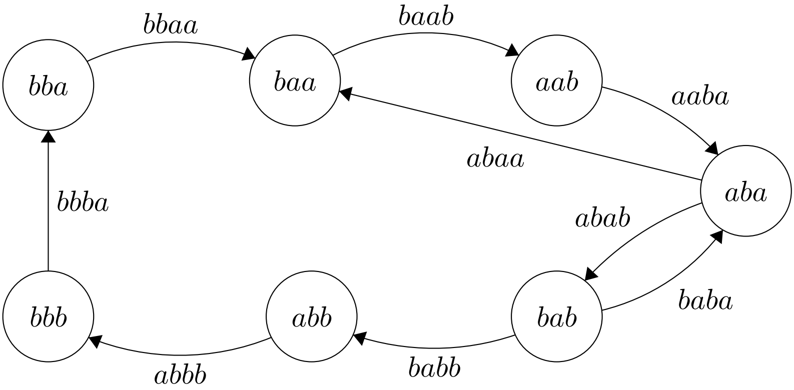
- Definitions : Let w be a finite word. For any integer i satisfying $1 \leq i \leq |w|$, let $L_w(i)$ denote the set of length-i factors of w and let $C_w(i)$ be the cardinality of $L_w(i)$.
- The i-th Rauzy graph of w is $\Gamma_w(i) = \{v_w(i), e_w(i)\}$ such that
  (i) $v_w(i) = L_w(i); e_w(i) = L_w(i+1)$,
  (ii)$e \in e_w(i+1)$ from v_1 to v_2 iff there exist $\alpha, \beta \in \mathrm{Alph}(w), y \in L_w(i-1)$ such that $v_1 = \alpha y, v_2 = y\beta, e = \alpha y \beta$.

Example

- Let us consider the 3-rd Rauzy graph of the word $w = bbaababaababbba$

Example

- Let us consider the 3-rd Rauzy graph of the word $w = bbaababaababbba$



Facts of Rauzy graphs

- For the graph $\Gamma_w(i)$,

$$|v_w(i)| = C_w(i);\ |e_w(i)| = C_w(i+1);$$

- The Rauzy graph $\Gamma_w(i)$ is weakly connected for all $1 \leq i \leq |w|$.
- The Rauzy graphs Contain cycles and circuits.
- The square factors in w are recognised with the help of circuits.

Facts of Rauzy graphs

- For the graph $\Gamma_w(i)$,

$$|v_w(i)| = C_w(i);\ |e_w(i)| = C_w(i+1);$$

- The Rauzy graph $\Gamma_w(i)$ is weakly connected for all $1 \leq i \leq |w|$.
- The Rauzy graphs Contain cycles and circuits.
- The square factors in w are recognised with the help of circuits.

Facts of Rauzy graphs

- For the graph $\Gamma_w(i)$,

$$|v_w(i)| = C_w(i);\ |e_w(i)| = C_w(i+1);$$

- The Rauzy graph $\Gamma_w(i)$ is weakly connected for all $1 \leq i \leq |w|$.
- The Rauzy graphs Contain cycles and circuits.
- The square factors in w are recognised with the help of circuits.

Facts of Rauzy graphs

- For the graph $\Gamma_w(i)$,

$$|v_w(i)| = C_w(i); \; |e_w(i)| = C_w(i+1);$$

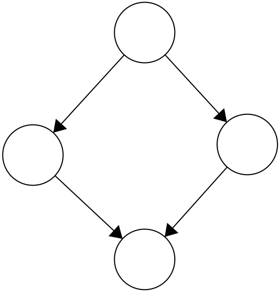
- The Rauzy graph $\Gamma_w(i)$ is weakly connected for all $1 \leq i \leq |w|$.
- The Rauzy graphs Contain cycles and circuits.
- The square factors in w are recognised with the help of circuits.

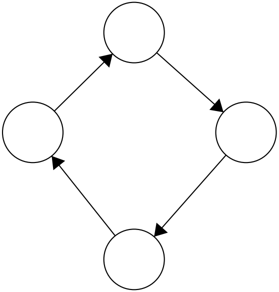
Basics on graph theory

Let $G(V,E)$ be an oriented graph, An elementary cycle is a closed chain of distinct vertices and distinct edges, and an elementary circuit is a closed path of distinct vertices and distinct edges.

Basics on graph theory

Let $G(V, E)$ be an oriented graph, An elementary cycle is a closed chain of distinct vertices and distinct edges, and an elementary circuit is a closed path of distinct vertices and distinct edges.





Basics of graph theory

Theorem (Berge(1973))

The number of the independent elementary cycles on a graph G with p connected components is given by the cyclomatic number

$$\chi(G) = e - v + p.$$

Corollary

For a given word w, the number of the independent elementary circuits on the graph $\Gamma_w(i)$ is bounded by

$$\chi(\Gamma_w(i)) = C_w(i+1) - C_w(i) + 1.$$

Basics of graph theory

Theorem (Berge(1973))

The number of the independent elementary cycles on a graph G with p connected components is given by the cyclomatic number

$$\chi(G) = e - v + p.$$

Corollary

For a given word w, the number of the independent elementary circuits on the graph $\Gamma_w(i)$ is bounded by

$$\chi(\Gamma_w(i)) = C_w(i+1) - C_w(i) + 1.$$

How the circuits help

Let $X = aab$.

If $w = ...a, a, b, a, a, b, ...$, then, on $\Gamma_w(3)$, there exists $C_1 = \{v_1, e_1\}$, such that v_1 is all conjugates of x.

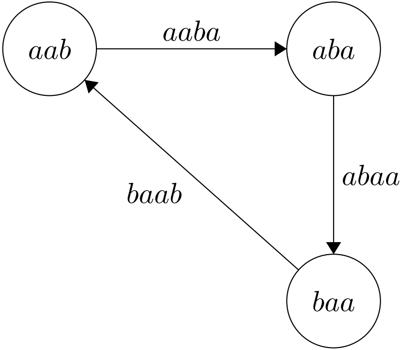
$$
\begin{tikzcd}
aab \arrow[rr, "aaba"] & & aba \arrow[d, "abaa"] \\
& & baa \arrow[llu, "baab"]
\end{tikzcd}
$$

How the circuits help

Let $X = aab$.
If $w = ...a, a, b, a, a, b, ...$, then, on $\Gamma_w(3)$, there exists $C_1 = \{v_1, e_1\}$, such that v_1 is all conjugates of x.

How the circuits help

Let $X = aab$.
If $w = ...{\color{red}a, a, b, a, a, b,} ...$, then, on $\Gamma_w(3)$, there exists $C_1 = \{v_1, e_1\}$, such that v_1 is all conjugates of x.

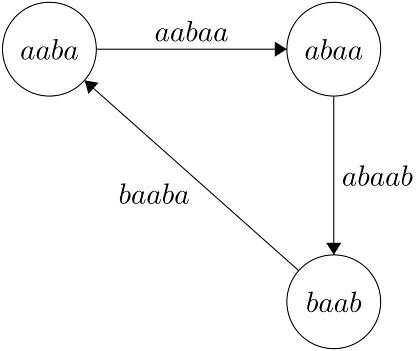


How the circuits help

If $w = ...a, a, b, a, a, b, ...a, b, a, a, b, a, ...$, then, on $\Gamma_w(4)$, there exists $C_2 = \{v_2, e_2\}$, such that $v_2 = e_1$.

How the circuits help

If $w = ...a, a, b, a, a, b, ...a, b, a, a, b, a, ...$, then, on $\Gamma_w(4)$, there exists $C_2 = \{v_2, e_2\}$, such that $v_2 = e_1$.

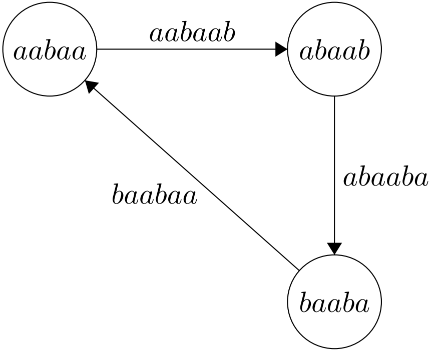


How the circuits help

If $w = \ldots a, a, b, a, a, b, \ldots a, b, a, a, b, a, a, \ldots$, then, on $\Gamma_w(5)$, there exists $C_3 = \{v_3, e_3\}$, such that $v_3 = e_2$.

How the circuits help

If $w = ...a, a, b, a, a, b, ...a, b, a, a, b, a, a, ...$, then, on $\Gamma_w(5)$, there exists $C_3 = \{v_3, e_3\}$, such that $v_3 = e_2$.

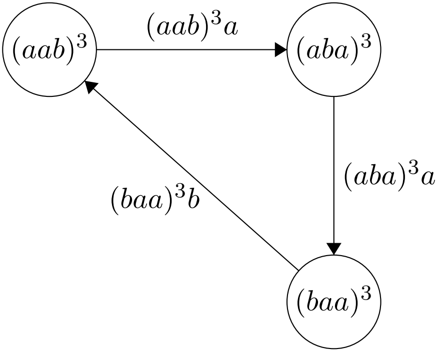


How the circuits help

If $w = ...(a, a, b)^4, ...$, then, on $\Gamma_w(9)$, there exists $C_4 = \{v_4, e_4\}$, such that v_4 is all conjugates of x^4.

How the circuits help

If $w = ...(a,a,b)^4, ...$, then, on $\Gamma_w(9)$, there exists $C_4 = \{v_4, e_4\}$, such that v_4 is all conjugates of x^4.

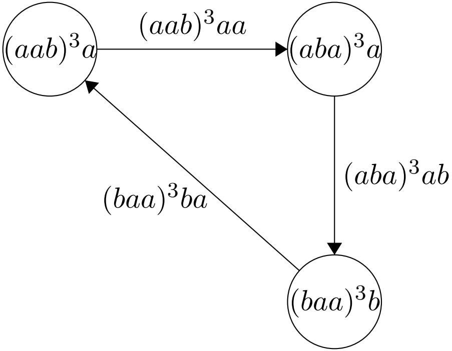


How the circuits help

If $w = ...(a, a, b)^4, a...$, then the word $(aba)^4$ can be associated with $C_5 = \{v_5, e_5\}$ on $\Gamma_w(10)$ such that $v_5 = e_4$.

How the circuits help

If $w = ...(a,a,b)^4, a...$, then the word $(aba)^4$ can be associated with $C_5 = \{v_5, e_5\}$ on $\Gamma_w(10)$ such that $v_5 = e_4$.



An injection

- There exists an injection from the set of nonempty square factors of w into its Rauzy graphs $\cup\Gamma_w(i)$:

$$\Phi : \quad Sq^*(w) \rightarrow \cup_{1\leq i\leq |w|}\Gamma_w(i).$$

- The images of Φ are small circuits.
- A circuit $C(V,E)$ on $\Gamma_w(i)$ is called small if it is an elementary circuit and $|V| = |E| \leq i$.

Proposition

All small circuits on the same Rauzy graph are independent.

An injection

- There exists an injection from the set of nonempty square factors of w into its Rauzy graphs $\cup\Gamma_w(i)$:

$$\Phi : \quad Sq^*(w) \to \cup_{1\leq i\leq |w|}\Gamma_w(i).$$

- The images of Φ are small circuits.
- A circuit $C(V, E)$ on $\Gamma_w(i)$ is called small if it is an elementary circuit and $|V| = |E| \leq i$.

Proposition

All small circuits on the same Rauzy graph are independent.

An injection

- There exists an injection from the set of nonempty square factors of w into its Rauzy graphs $\cup\Gamma_w(i)$:

$$\Phi : \quad Sq^*(w) \to \cup_{1\leq i\leq |w|}\Gamma_w(i).$$

- The images of Φ are small circuits.
- A circuit $C(V, E)$ on $\Gamma_w(i)$ is called small if it is an elementary circuit and $|V| = |E| \leq i$.

Proposition

All small circuits on the same Rauzy graph are independent.

An injection

- There exists an injection from the set of nonempty square factors of w into its Rauzy graphs $\cup\Gamma_w(i)$:

$$\Phi : \quad Sq^*(w) \to \cup_{1\leq i\leq |w|}\Gamma_w(i).$$

- The images of Φ are small circuits.
- A circuit $C(V,E)$ on $\Gamma_w(i)$ is called small if it is an elementary circuit and $|V| = |E| \leq i$.

Proposition

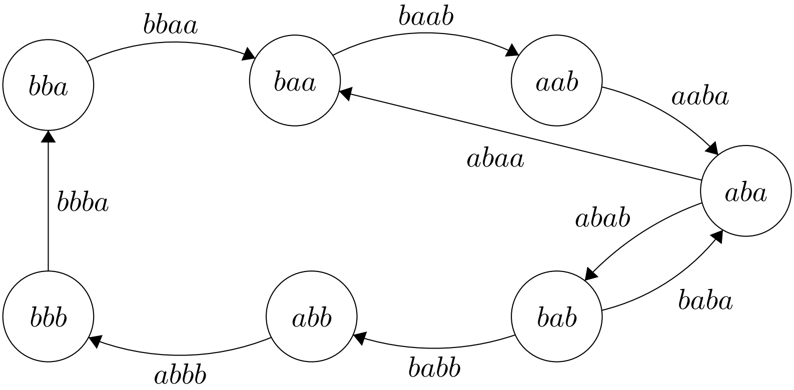
All small circuits on the same Rauzy graph are independent.

Example

- Let us consider the 3-rd Rauzy graph of the word $w = bbaababaababbba$

Example

- Let us consider the 3-rd Rauzy graph of the word $w = bbaababaababbba$



End of the proof

For a given word w and an integer i satisfying $1 \leq i \leq |w|$,

$$|Im(\Phi) \cap \Gamma_w(i)| \leq \chi(\Gamma_w(i)) = C_w(i+1) - C_w(i) + 1.$$

Thus,

$$\begin{aligned} S^*(w) &= \sum_{1 \leq i \leq |w|} |Im(\Phi) \cap \Gamma_w(i)| \\ &\leq \sum_{1 \leq i \leq |w|} C_w(i+1) - C_w(i) + 1 \\ &\leq |w| - |\mathrm{Alph}(w)|. \end{aligned}$$

End of the proof

For a given word w and an integer i satisfying $1 \leq i \leq |w|$,

$$|Im(\Phi) \cap \Gamma_w(i)| \leq \chi(\Gamma_w(i)) = C_w(i+1) - C_w(i) + 1.$$

Thus,

$$\begin{aligned} S^*(w) &= \sum_{1 \leq i \leq |w|} |Im(\Phi) \cap \Gamma_w(i)| \\ &\leq \sum_{1 \leq i \leq |w|} C_w(i+1) - C_w(i) + 1 \\ &\leq |w| - |\mathrm{Alph}(w)|. \end{aligned}$$

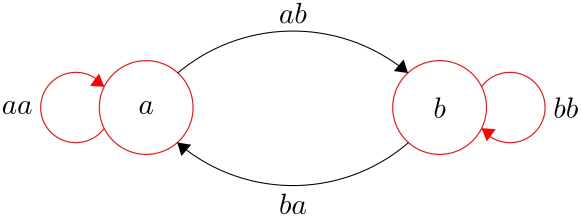
Example

$w = \color{red}{bbaa}\color{black}{babaababbba}$
$\Gamma_w(1)$

Example

$w = \textcolor{red}{bbaa}babaababbba$

$\Gamma_w(1)$



Example

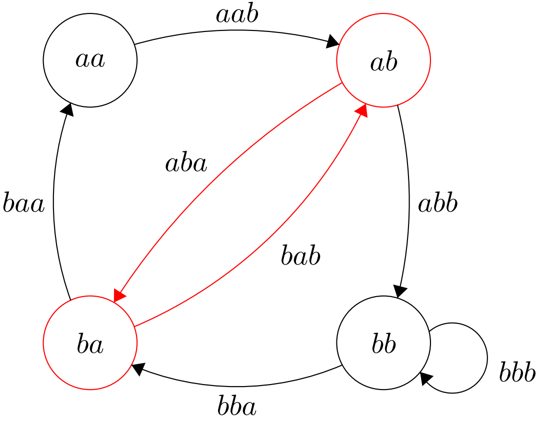
$w = bba\color{red}{abab}\color{black}{aababbba}$

$\Gamma_w(2)$

Example

$w = bba\color{red}{abab}aababbba$

$\Gamma_w(2)$



Example

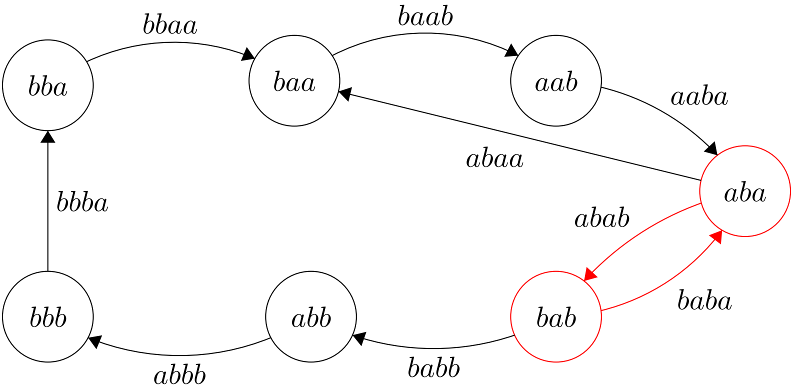
Let us consider the 3-rd Rauzy graph of the word
$w = bbaa\textcolor{red}{baba}ababbba$
$\Gamma_w(3)$

Example

Let us consider the 3-rd Rauzy graph of the word

$w = bbaababaababbba$

$\Gamma_w(3)$

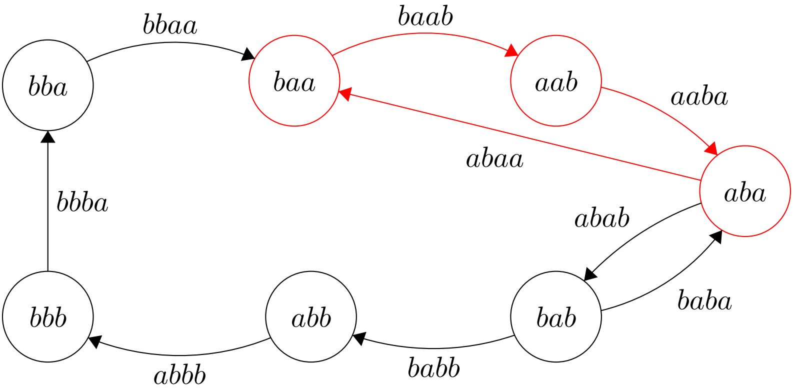


Example

Let us consider the 3-rd Rauzy graph of the word
$w = bbaab\textcolor{red}{abaaba}bbba$
$\Gamma_w(3)$

Example

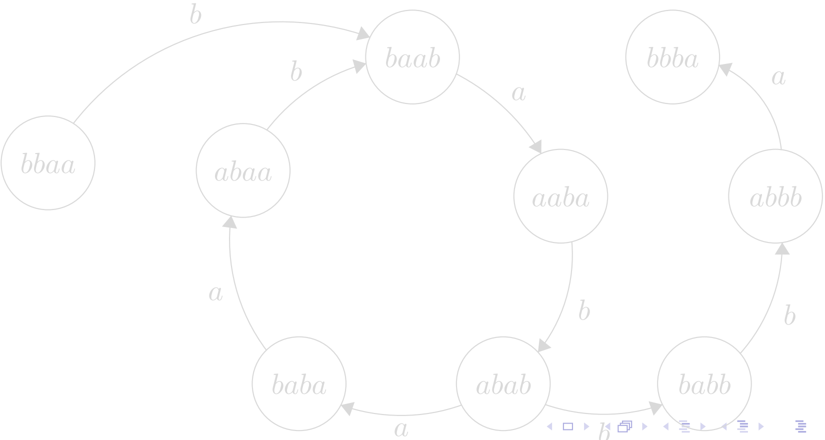
Let us consider the 3-rd Rauzy graph of the word
$w = bbaababaababbbba$
$\Gamma_w(3)$



Example

$w = bbaababaababbba$

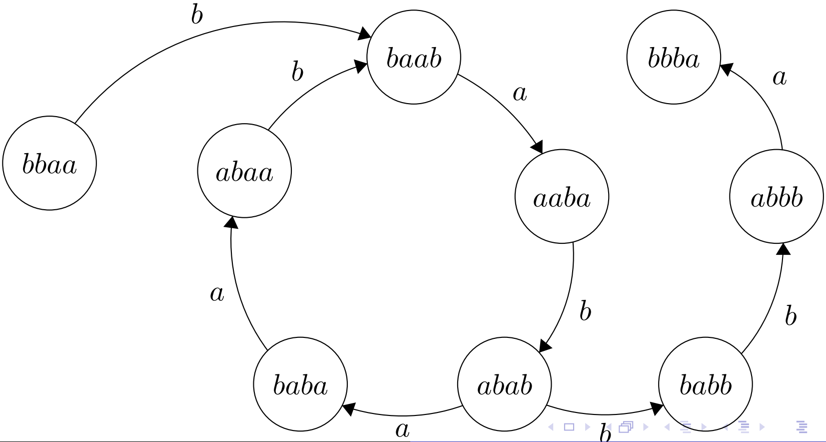
$\Gamma_w(4)$



Example

$w = bbaababaababbba$

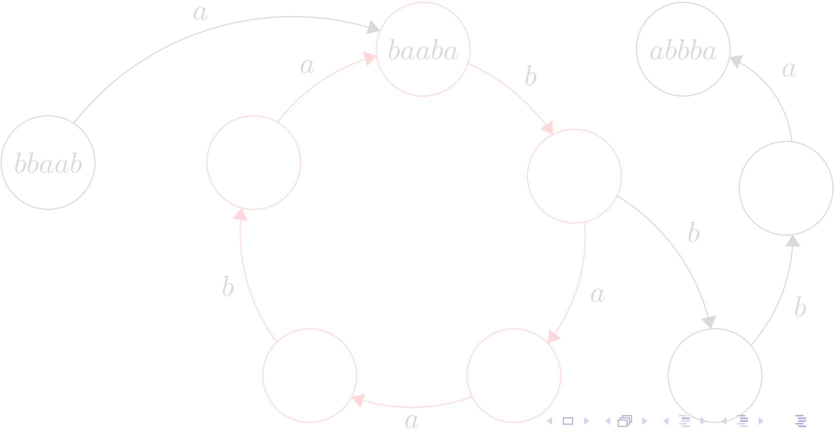
$\Gamma_w(4)$



Example

$w = b\color{red}{baababaaba}\color{black}{bbba}$

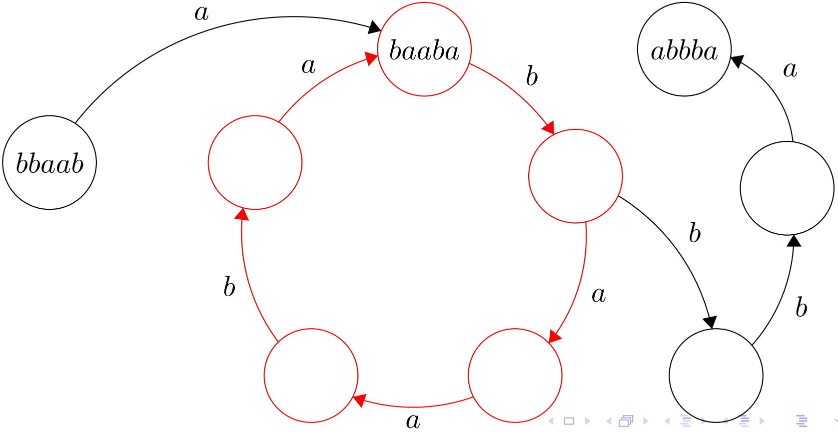
$\Gamma_w(5)$



Example

$w = b\color{red}{baababaaba}\color{black}{bbba}$

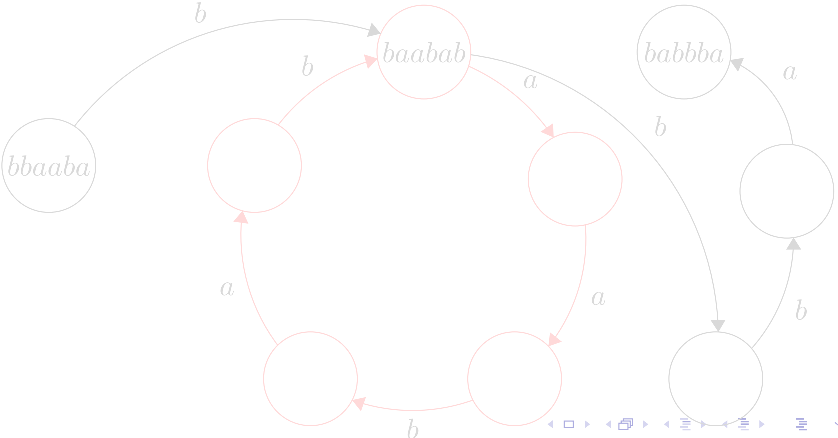
$\Gamma_w(5)$



Example

$w = bb\textcolor{red}{aababaabab}bba$

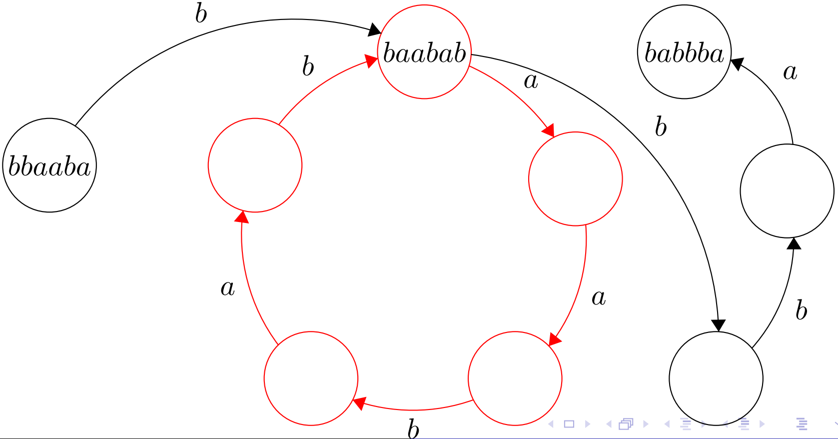
$\Gamma_w(6)$



Example

$w = bb\textcolor{red}{aababaabab}bba$

$\Gamma_w(6)$



(Non)-sharpness : reason 1

Theorem

The upper bound $s^*(w) \leq |w| - |\mathrm{Alph}(w)|$ is not sharp.

- Reason 1 : the gaps.
- Let $x = aab$ and $w = ...x^4....$ From our previous construction, there exists a function Φ such that $(aab)^2 \to \Gamma_w(3), (aba)^2 \to \Gamma_w(4), (baa)^2 \to \Gamma_w(5), (aab)^4 \to \Gamma_w(9)$.

(Non)-sharpness : reason 1

Theorem

The upper bound $s^*(w) \leq |w| - |\mathrm{Alph}(w)|$ is not sharp.

- Reason 1 : the gaps.
- Let $x = aab$ and $w = ...x^4....$ From our previous construction, there exists a function Φ such that $(aab)^2 \to \Gamma_w(3), (aba)^2 \to \Gamma_w(4), (baa)^2 \to \Gamma_w(5), (aab)^4 \to \Gamma_w(9)$.

(Non)-sharpness : reason 1

Theorem

The upper bound $s^*(w) \leq |w| - |\mathrm{Alph}(w)|$ is not sharp.

- Reason 1 : the gaps.
- Let $x = aab$ and $w = ...x^4....$ From our previous construction, there exists a function Φ such that $(aab)^2 \to \Gamma_w(3), (aba)^2 \to \Gamma_w(4), (baa)^2 \to \Gamma_w(5), (aab)^4 \to \Gamma_w(9)$.

(Non)-sharpness : reason 1

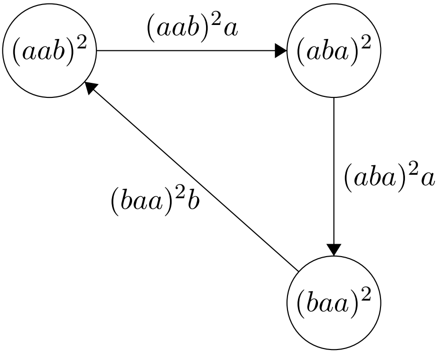
But there also exist similar circuits on $\Gamma_w(6), \Gamma_w(7), \Gamma_w(8)$. On $\Gamma_w(6)$:

(Non)-sharpness : reason 1

But there also exist similar circuits on $\Gamma_w(6), \Gamma_w(7), \Gamma_w(8)$. On $\Gamma_w(6)$:

(Non)-sharpness : reason 1

But there also exist similar circuits on $\Gamma_w(6), \Gamma_w(7), \Gamma_w(8)$. On $\Gamma_w(6)$:

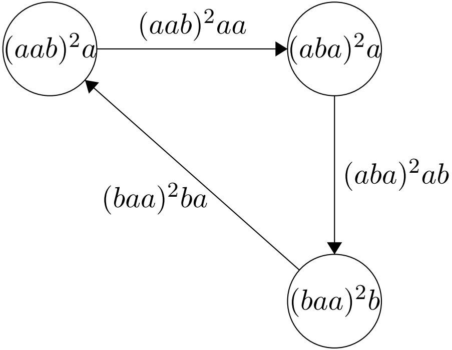


(Non)-sharpness : reason 1

On $\Gamma_w(7)$:

(Non)-sharpness : reason 1

On $\Gamma_w(7)$:

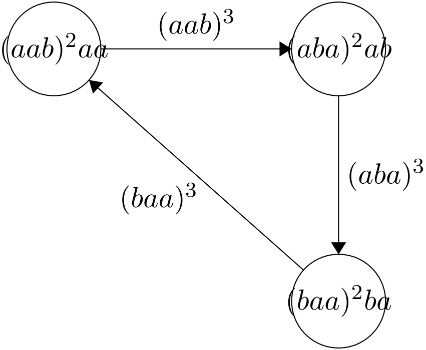


(Non)-sharpness : reason 1

On $\Gamma_w(8)$:

(Non)-sharpness : reason 1

On $\Gamma_w(8)$:



A generalisation

- A power is a word of the from $u^k = \underbrace{uu...u}_{k \text{ times}}$. It is also called a k-power, and k is its exponent.
- For a finite word w, let $M^*(w)$ denote the set of nonempty powers of exponent at least 2 in w.
- There exists an injection from $M^*(w)$ into its Rauzy graphs $\cup\Gamma_w(i)$:

$$\Psi : \quad M^*(w) \to \cup_{1\leq i\leq |w|}\Gamma_w(i).$$

A generalisation

- A power is a word of the from $u^k = \underbrace{uu...u}_{k \text{ times}}$. It is also called a k-power, and k is its exponent.
- For a finite word w, let $M^*(w)$ denote the set of nonempty powers of exponent at least 2 in w.
- There exists an injection from $M^*(w)$ into its Rauzy graphs $\cup\Gamma_w(i)$:

$$\Psi : \quad M^*(w) \to \cup_{1\leq i\leq |w|}\Gamma_w(i).$$

A generalisation

- A power is a word of the from $u^k = \underbrace{uu...u}_{k \text{ times}}$. It is also called a k-power, and k is its exponent.
- For a finite word w, let $M^*(w)$ denote the set of nonempty powers of exponent at least 2 in w.
- There exists an injection from $M^*(w)$ into its Rauzy graphs $\cup\Gamma_w(i)$:

$$\Psi : \quad M^*(w) \to \cup_{1\leq i\leq |w|}\Gamma_w(i).$$

A generalisation

Theorem (Li,Pachocki and Radoszewski 2022)

For every finite word w, let $m^*(w)$ denote the number of distinct nonempty powers of exponent at least 2 in w, let $m^*_k(w)$ denote the number of distinct nonempty k-powers in w, then one has

$$m^*(w) \leq |w| - |\mathrm{Alph}(w)|;$$

Moreover, for any integer $k \geq 2$,

$$m^*_k(w) \leq \frac{|w| - |\mathrm{Alph}(w)|}{k-1}.$$

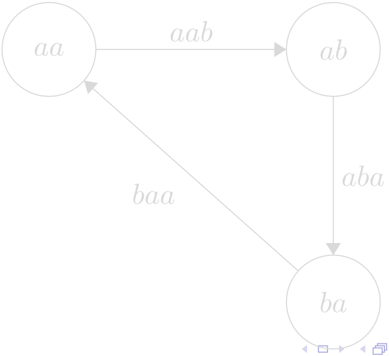
(Non)-sharpness : reason 2

- The Reason 1 is not critical : The w may not contain any other powers of exponent larger than 2.
- Reason 2 : Existence of other independent circuits.
- Let $x = aab$ and let $w = ...x^2...$, then there exists the following circuit on $\Gamma_w(2)$:

$$
\begin{tikzcd}
aa \arrow[rr, "aab"] & & ab \arrow[d, "aba"] \\
& & ba \arrow[llu, "baa"]
\end{tikzcd}
$$

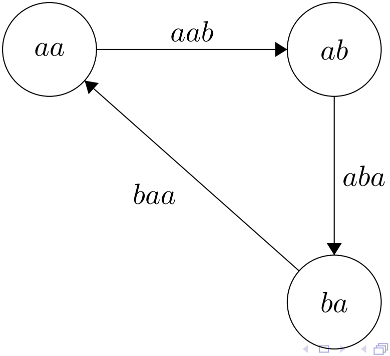
(Non)-sharpness : reason 2

- The Reason 1 is not critical : The w may not contain any other powers of exponent larger than 2.
- Reason 2 : Existence of other independent circuits.
- Let $x = aab$ and let $w = ...x^2...$, then there exists the following circuit on $\Gamma_w(2)$:



(Non)-sharpness : reason 2

- The Reason 1 is not critical : The w may not contain any other powers of exponent larger than 2.
- Reason 2 : Existence of other independent circuits.
- Let $x = aab$ and let $w = ...x^2...$, then there exists the following circuit on $\Gamma_w(2)$:



Primitive circuits

- Let x be a primitive word and let $[x]$ be the conjugate class of x.
- If there exists a word w such that $x^2 \in \mathrm{Fact}(w)$ and if $|x| \geq 2$, then there exists a circuit on $\Gamma_w(|x|-1)$ such that its edge set is $[x]$. This circuit is called a primitive circuit.

Proposition (Li, 2022)

The primitive circuits are independent with small circuits.

Primitive circuits

- Let x be a primitive word and let $[x]$ be the conjugate class of x.
- If there exists a word w such that $x^2 \in \mathrm{Fact}(w)$ and if $|x| \geq 2$, then there exists a circuit on $\Gamma_w(|x|-1)$ such that its edge set is $[x]$. This circuit is called a primitive circuit.

Proposition (Li, 2022)

The primitive circuits are independent with small circuits.

Primitive circuits

- Let x be a primitive word and let $[x]$ be the conjugate class of x.
- If there exists a word w such that $x^2 \in \mathrm{Fact}(w)$ and if $|x| \geq 2$, then there exists a circuit on $\Gamma_w(|x|-1)$ such that its edge set is $[x]$. This circuit is called a primitive circuit.

Proposition (Li, 2022)

The primitive circuits are independent with small circuits.

(Non)-sharpness : reason 3

- Reason 3 : Simulation.
- A248958 : maximum number of distinct nonempty squares in a binary string of length n.

$$B(n) = 0, 1, 1, 2, 2, 3, 3, 4, 5, 6, 7, 7, 8, 9, 10, 11, 12, 12, 13, 13, 14, 15, 16, 17, 18, 19, 20, 20, 21, 22, 23, 23, 24, 25$$

Conjecture (Brlek and Li, 2022)

Let $MS(n)$ denote the maximum number of distinct nonempty squares in a word of length n, then

$$MS(n) \leq \lceil n + 1 - \sqrt{n} - \log_2 \sqrt{n} \rceil = Up(n).$$

(Non)-sharpness : reason 3

- Reason 3 : Simulation.
- A248958 : maximum number of distinct nonempty squares in a binary string of length n.

$$B(n) = 0, 1, 1, 2, 2, 3, 3, 4, 5, 6, 7, 7, 8, 9, 10, 11, 12, 12, 13, 13, 14, 15, 16, 17, 18, 19, 20, 20, 21, 22, 23, 23, 24, 25$$

Conjecture (Brlek and Li, 2022)

Let $MS(n)$ denote the maximum number of distinct nonempty squares in a word of length n, then

$$MS(n) \leq \lceil n + 1 - \sqrt{n} - \log_2 \sqrt{n} \rceil = Up(n).$$

(Non)-sharpness : reason 3

- Reason 3 : Simulation.
- A248958 : maximum number of distinct nonempty squares in a binary string of length n.

$$B(n) = 0, 1, 1, 2, 2, 3, 3, 4, 5, 6, 7, 7, 8, 9, 10, 11, 12, 12, 13, 13, 14, 15, 16, 17, 18, 19, 20, 20, 21, 22, 23, 23, 24, 25$$

Conjecture (Brlek and Li, 2022)

Let $MS(n)$ denote the maximum number of distinct nonempty squares in a word of length n, then

$$MS(n) \leq \lceil n + 1 - \sqrt{n} - \log_2 \sqrt{n} \rceil = Up(n).$$

Some remarks and open questions

- $$Up(n) - B(n) = 1,1,1,0,1,1,1,1,1,1,0,1,1,1,1,0,0,1,1,2,2,2,1,1,1,1,1,2,2,2,1,2,2,2$$
- If there exists some word w satisfying $s(w) = MS(|w|)$, then w is a binary word ?
- $MS(n+1) - MS(n) \leq 1$?
- How many independent circuits can a Rauzy graph contain ?

Some remarks and open questions

- $$Up(n) - B(n) = 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 2, 2, 2, 1, 1, 1, 1, 1, 2, 2, 2, 1, 2, 2, 2$$
- If there exists some word w satisfying $s(w) = MS(|w|)$, then w is a binary word ?
- $MS(n+1) - MS(n) \leq 1$?
- How many independent circuits can a Rauzy graph contain ?

Some remarks and open questions

- $$Up(n) - B(n) = 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 2, 2, 2, 1, 1, 1, 1, 1, 2, 2, 2, 1, 2, 2, 2$$
- If there exists some word w satisfying $s(w) = MS(|w|)$, then w is a binary word ?
- $MS(n+1) - MS(n) \leq 1$?
- How many independent circuits can a Rauzy graph contain ?

Some remarks and open questions

- $$Up(n) - B(n) = 1,1,1,0,1,1,1,1,1,1,0,1,1,1,1,0,0,1,1,2,2,\\ 2,1,1,1,1,1,2,2,2,1,2,2,2$$
- If there exists some word w satisfying $s(w) = MS(|w|)$, then w is a binary word ?
- $MS(n+1) - MS(n) \leq 1$?
- How many independent circuits can a Rauzy graph contain ?

Motivation

Theorem (Brlek, Li (2022))

Any finite word w, if we let $|\mathrm{Alph}(w)|$ denote the number of distinct letters in w, one has

$$s^*(w) \leq |w| - |\mathrm{Alph}(w)|.$$

Motivation

Let w be a finite word.

- The palindromic defect of w is defined to be

$$D_p(w) = |w| + 1 - p(w),$$

- The σ-palindromic defect of w is defined to be

$$D_{p_\sigma}(w) = |w| + 1 - p_\sigma(w).$$

Motivation

Let w be a finite word.

- The palindromic defect of w is defined to be

$$D_p(w) = |w| + 1 - p(w),$$

- The σ-palindromic defect of w is defined to be

$$D_{p_\sigma}(w) = |w| + 1 - p_\sigma(w).$$

BR-identity

Theorem

For any finite word w, one has

$$2D_p(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 2 - p_w(i+1) - p_w(i),$$

where $p_w(i)$ is the number of distinct length-i palindromes in w.(Brlek, Reutenauer 2011)

$$2D_{p_\sigma}(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 2 - p_{\sigma w}(i+1) - p_{\sigma w}(i),$$

where $p_{\sigma w}(i)$ is the number of distinct length-i σ-palindromes in w.(Reutenauer and Fest 2013)

Square defect

- Let w be a finite word, the square defect of w is defined to be

$$D_s(w) = |w| + 1 - s(w),$$

Theorem (Brlek, Dolce, Vandomme (2018))

For any finite word w, one has

$$2D_s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 2 - s_w(i+1) - s_w(i),$$

where $s_w(i)$ is the number of distinct length-i squares in w.

Square defect

- Let w be a finite word, the square defect of w is defined to be
$$D_s(w) = |w| + 1 - s(w),$$

Theorem (Brlek, Dolce, Vandomme (2018))

For any finite word w, one has

$$2D_s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 2 - s_w(i+1) - s_w(i),$$

where $s_w(i)$ is the number of distinct length-i squares in w.

Alternative BR-identity

$$2D_s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 2 - s_w(i+1) - s_w(i)$$

$$\iff s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 1 - D_s(w).$$

$$s(w) \leq \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 1 \iff s(w) \leq |w| + 1 - |\mathrm{Alph}(w)|.$$

Fraenkel-Simpson conjecture $\iff$ $D_s(w) \geq 0$.

Alternative BR-identity

$$2D_s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 2 - s_w(i+1) - s_w(i)$$

$$\iff s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 1 - D_s(w).$$

$$s(w) \leq \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 1 \iff s(w) \leq |w| + 1 - |\mathrm{Alph}(w)|.$$

Fraenkel-Simpson conjecture $\iff$ $D_s(w) \geq 0$.

Alternative BR-identity

$$2D_s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 2 - s_w(i+1) - s_w(i)$$

$$\iff s(w) = \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 1 - D_s(w).$$

$$s(w) \leq \sum_{i=0}^{|w|} C_w(i+1) - C_w(i) + 1 \iff s(w) \leq |w| + 1 - |\mathrm{Alph}(w)|.$$

Fraenkel-Simpson conjecture $\iff$ $D_s(w) \geq 0.$

Square defect for infinite words

Let w be an infinite word, its square defect is defined to be

$$D_s(w) = \sup \{D_s(p) | p \in \mathrm{Fact}(w)\}.$$

Theorem (Brlek, Dolce, Vandomme (2018))

The square defect of any periodic word or strict standard episturmian word is infinite.
The square BR-identity holds for any infinite periodic word and strict standard episturmian.

Square defect for infinite words

Let w be an infinite word, its square defect is defined to be

$$D_s(w) = \sup \{D_s(p) | p \in \text{Fact}(w)\} .$$

Theorem (Brlek, Dolce, Vandomme (2018))

The square defect of any periodic word or strict standard episturmian word is infinite.
The square BR-identity holds for any infinite periodic word and strict standard episturmian.

Conjectures

Conjecture

The square defect of any infinite word is infinite.

Conjecture

The square BR-identity holds for any infinite word.

Conjectures

Conjecture

The square defect of any infinite word is infinite.

Conjecture

The square BR-identity holds for any infinite word.

Thank you for your attention !