Repetition Thresholds in Graphs

Borut Lužar

Faculty of Information Studies, Novo mesto, Slovenia

borut.luzar@gmail.com
http://luzar.fis.unm.si

joint work with Pascal Ochem & Alex Pinlou

Grafy 2019 - June 7, 2019

• Given an alphabet of k letters

$$\mathbb{A} = \{a_1, a_2, \ldots, a_k\},\$$

a word of length n over \mathbb{A} ,

$$w = \ell_1 \ell_2 \cdots \ell_n,$$

is a sequence of letters from \mathbb{A} , i.e., $\ell_i \in \mathbb{A}$, for every $1 \leq i \leq n$.

• Given an alphabet of k letters

$$\mathbb{A} = \{a_1, a_2, \ldots, a_k\},\$$

a word of length n over \mathbb{A} ,

$$w = \ell_1 \ell_2 \cdots \ell_n,$$

is a sequence of letters from \mathbb{A} ,

- i.e., $\ell_i \in \mathbb{A}$, for every $1 \leq i \leq n$.
- We are interested in consecutive repetitions of subwords or their parts;

A subword or a factor of a word w is a sequence of consecutive letters in w.

A subword or a factor of a word w is a sequence of consecutive letters in w.

ok is a factor of kokos

A subword or a factor of a word w is a sequence of consecutive letters in w.

ok is a factor of kokos

A repetition in a word w is a factor consisting of two identical consecutive factors

■ A subword or a factor of a word *w* is a sequence of consecutive letters in *w*.

ok is a factor of kokos

A repetition in a word w is a factor consisting of two identical consecutive factors

kokos banana ananas

■ A subword or a factor of a word *w* is a sequence of consecutive letters in *w*.

ok is a factor of kokos

A repetition in a word w is a factor consisting of two identical consecutive factors

kokos banana ananas

A word is non-repetitive if it does not contain a repetition;

 Thue [18] showed there exist infinite non-repetitive words on 3 letters;

- Thue [18] showed there exist infinite non-repetitive words on 3 letters;
- Various constructions exist;

- Thue [18] showed there exist infinite non-repetitive words on 3 letters;
- Various constructions exist;
- Thue-Morse (0, 1) generator:
 - start with $w_0 = 1$;

- Thue [18] showed there exist infinite non-repetitive words on 3 letters;
- Various constructions exist;
- Thue-Morse (0,1) generator:
 - start with $w_0 = 1$;
 - define morphism m as m(0) = 01 and m(1) = 10;

- Thue [18] showed there exist infinite non-repetitive words on 3 letters;
- Various constructions exist;
- Thue-Morse (0, 1) generator:
 - start with $w_0 = 1$;
 - define morphism m as m(0) = 01 and m(1) = 10;
 - define $w_i = m(w_{i-1})$;

- Thue [18] showed there exist infinite non-repetitive words on 3 letters;
- Various constructions exist;
- Thue-Morse (0, 1) generator:
 - start with $w_0 = 1$;
 - define morphism m as m(0) = 01 and m(1) = 10;
 - define $w_i = m(w_{i-1})$;
 - from a word w_i we construct a word x_i as a sequence of numbers of zeros between each pair of consecutive ones in w_i;

- Thue [18] showed there exist infinite non-repetitive words on 3 letters;
- Various constructions exist;
- Thue-Morse (0, 1) generator:
 - start with $w_0 = 1$;
 - define morphism m as m(0) = 01 and m(1) = 10;
 - define $w_i = m(w_{i-1})$;
 - from a word w_i we construct a word x_i as a sequence of numbers of zeros between each pair of consecutive ones in w_i;
- Example: $w_6 = 1001011001101001011010011010010110$

$$x_6 = 210201210120210$$

What if there are several equal consecutive factors?

- What if there are several equal consecutive factors?
- Having an alphabet with only two letters

$$\mathbb{A} = \{a, b\}$$

it is impossible to construct a long non-repetitive word

- What if there are several equal consecutive factors?
- Having an alphabet with only two letters

$$\mathbb{A} = \{a, b\}$$

it is impossible to construct a long non-repetitive word *a*

- What if there are several equal consecutive factors?
- Having an alphabet with only two letters

$$\mathbb{A} = \{a, b\}$$

it is impossible to construct a long non-repetitive word a ab

- What if there are several equal consecutive factors?
- Having an alphabet with only two letters

$$\mathbb{A} = \{a, b\}$$

it is impossible to construct a long non-repetitive word *a ab ab ab*

- What if there are several equal consecutive factors?
- Having an alphabet with only two letters

$$\mathbb{A} = \{a, b\}$$

it is impossible to construct a long non-repetitive word *a ab aba aba aba*

- What if there are several equal consecutive factors?
- Having an alphabet with only two letters

$$\mathbb{A} = \{a, b\}$$

it is impossible to construct a long non-repetitive word

i a ab aba aba?

 Thue proved that there is an arbitrarily long word on two letters without three equal consecutive factors;

What if we do not require repetition of whole factors?

What if we do not require repetition of whole factors?

abc abc ab

• What if we do not require repetition of whole factors? *abc abc ab abc* is repeated $2 + \frac{2}{3} = \frac{8}{3}$ times;

What if we do not require repetition of whole factors?

abc abc ab abc is repeated $2 + \frac{2}{3} = \frac{8}{3}$ times;

How low repetitiveness can be achieved with alphabets on k letters?

A generalized definition of a repetition:

A prefix of a word w = w₁...w_r is a word x = w₁...w_s, for some s ≤ r;

A generalized definition of a repetition:

A prefix of a word w = w₁...w_r is a word x = w₁...w_s, for some s ≤ r;

w = banana x = ban

A generalized definition of a repetition:

- A prefix of a word w = w₁...w_r is a word x = w₁...w_s, for some s ≤ r;
 - w = banana x = ban
- A repetition in a word w is a pair of words p (the period) and e (the excess) such that pe is a factor of w, p is non-empty, and e is a prefix of pe;

A generalized definition of a repetition:

A prefix of a word w = w₁...w_r is a word x = w₁...w_s, for some s ≤ r;

w = banana x = ban

A repetition in a word w is a pair of words p (the period) and e (the excess) such that pe is a factor of w, p is non-empty, and e is a prefix of pe;

w = banana pe = anana p = an e = ana

A generalized definition of a repetition:

A prefix of a word w = w₁...w_r is a word x = w₁...w_s, for some s ≤ r;

w = banana x = ban

A repetition in a word w is a pair of words p (the period) and e (the excess) such that pe is a factor of w, p is non-empty, and e is a prefix of pe;

w = banana pe = anana p = an e = ana

■ The exponent of a repetition *pe* is

$$\exp(pe) = \frac{|pe|}{|p|}$$

A generalized definition of a repetition:

A prefix of a word w = w₁...w_r is a word x = w₁...w_s, for some s ≤ r;

w = banana x = ban

A repetition in a word w is a pair of words p (the period) and e (the excess) such that pe is a factor of w, p is non-empty, and e is a prefix of pe;

w = banana pe = anana p = an e = ana

The exponent of a repetition pe is

$$\exp(pe) = \frac{|pe|}{|p|}$$

 $\exp(anana) = \frac{5}{2}$

• A β -repetition is a repetition of exponent β ;

- A β -repetition is a repetition of exponent β ;
- A word is α^+ -free if it contains no β -repetition such that $\beta > \alpha$;

- A β -repetition is a repetition of exponent β ;
- A word is α^+ -free if it contains no β -repetition such that $\beta > \alpha$;
- A word is α -free if it contains no β -repetition such that $\beta \geq \alpha$;

- A β -repetition is a repetition of exponent β ;
- A word is α^+ -free if it contains no β -repetition such that $\beta > \alpha$;
- A word is α -free if it contains no β -repetition such that $\beta \geq \alpha$;
- For k ≥ 2, the repetition threshold RT(k) for k letters is the smallest α such that there exists an infinite α⁺-free word over a k-letter alphabet;
Repetition Thresholds

- A β -repetition is a repetition of exponent β ;
- A word is α^+ -free if it contains no β -repetition such that $\beta > \alpha$;
- A word is α -free if it contains no β -repetition such that $\beta \geq \alpha$;
- For k ≥ 2, the repetition threshold RT(k) for k letters is the smallest α such that there exists an infinite α⁺-free word over a k-letter alphabet;
- The notion of repetition thresholds was initiated by Dejean [11] in 1972;

Repetition Thresholds - Results

Repetition thresholds for words are completely solved;

Repetition Thresholds - Results

Repetition thresholds for words are completely solved;

Theorem 1

(i)
$$\operatorname{RT}(2) = 2$$
 [11];
(ii) $\operatorname{RT}(3) = \frac{7}{4}$ [11];
(iii) $\operatorname{RT}(4) = \frac{7}{5}$ [16];
(iv) $\operatorname{RT}(k) = \frac{k}{k-1}$, for $k \ge 5$ [3, 8, 9, 10, 13, 14, 16, 17].

 Study of non-repetitiveness has been generalized to graphs by Currie [4, 5] and Alon et al. [2];

- Study of non-repetitiveness has been generalized to graphs by Currie [4, 5] and Alon et al. [2];
- For a k-vertex coloring of a graph, a sequence of colors on a non-intersecting path is called a factor;

- Study of non-repetitiveness has been generalized to graphs by Currie [4, 5] and Alon et al. [2];
- For a k-vertex coloring of a graph, a sequence of colors on a non-intersecting path is called a factor;
- A vertex coloring is said to be α^+ -free (resp. α -free) if every factor is α^+ -free (resp. α -free);

■ For a *k*-vertex colored graph *G*, the repetition threshold is

 $\operatorname{RT}(k,G) = \inf_{k \text{-coloring } c} \sup \left\{ \exp(w) \, | \, w \text{ is a factor in } c \right\} \, .$

For a *k*-vertex colored graph *G*, the repetition threshold is

$$\operatorname{RT}(k,G) = \inf_{k \operatorname{-coloring}} \sup_{c} \left\{ \exp(w) \, | \, w \text{ is a factor in } c \right\} \,.$$

The repetition threshold over a whole class of graphs G is defined as

$$\operatorname{RT}(k,\mathcal{G}) = \sup_{G\in\mathcal{G}}\operatorname{RT}(k,G).$$

 For the class of paths *P* the repetition thresholds are known, since

$$\operatorname{RT}(k, \mathcal{P}) = \operatorname{RT}(k);$$

Repetition Thresholds of Cycles - \mathcal{C}

Theorem 2

(i)
$$\operatorname{RT}(2, \mathcal{C}) = \frac{5}{2} [1];$$

(ii) $\operatorname{RT}(3, \mathcal{C}) = 2 [6];$
(iii) $\operatorname{RT}(4, \mathcal{C}) = \frac{3}{2} [7];$
(iv) $\operatorname{RT}(5, \mathcal{C}) = \frac{4}{3} [7];$
(v) $\operatorname{RT}(k, \mathcal{C}) = 1 + \frac{1}{\lceil \frac{k}{2} \rceil},$ for $k \ge 6$ [12].

(*i*)
$$\operatorname{RT}(2, \mathcal{T}) = \frac{7}{2};$$

(*ii*) $\operatorname{RT}(3, \mathcal{T}) = 3;$
(*iii*) $\operatorname{RT}(k, \mathcal{T}) = \frac{3}{2}, \text{ for } k \ge 4.$

(*i*) RT(2,
$$T$$
) = $\frac{7}{2}$;
(*ii*) RT(3, T) = 3;
(*iii*) RT(k, T) = $\frac{3}{2}$, for $k \ge 4$.



(*i*) RT(2,
$$T$$
) = $\frac{7}{2}$;
(*ii*) RT(3, T) = 3;
(*iii*) RT(k, T) = $\frac{3}{2}$, for $k \ge 4$.



(*i*) RT(2,
$$T$$
) = $\frac{7}{2}$;
(*ii*) RT(3, T) = 3;
(*iii*) RT(k, T) = $\frac{3}{2}$, for $k \ge 4$.



(*i*) RT(2,
$$T$$
) = $\frac{7}{2}$;
(*ii*) RT(3, T) = 3;
(*iii*) RT(k, T) = $\frac{3}{2}$, for $k \ge 4$.



(*i*) RT(2,
$$T$$
) = $\frac{7}{2}$;
(*ii*) RT(3, T) = 3;
(*iii*) RT(k, T) = $\frac{3}{2}$, for $k \ge 4$.



(*i*) RT(2,
$$T$$
) = $\frac{7}{2}$;
(*ii*) RT(3, T) = 3;
(*iii*) RT(k, T) = $\frac{3}{2}$, for $k \ge 4$.



Repetition Thresholds of Subdivisions - \mathcal{S}

Theorem 4 ([15])

(*i*) RT(2, S) =
$$\frac{7}{3}$$
;
(*ii*) RT(3, S) = $\frac{7}{4}$;
(*iii*) RT(k, S) = $\frac{3}{2}$, for $k \ge 4$.

A subdivision of a graph G is a graph obtained from G by a sequence of edge subdivisions. By a graph subdivision, we always mean a "large enough" subdivision.

Repetition Thresholds of Caterpillars - \mathcal{CP}

 A caterpillar is a tree such that the graph induced by the vertices of degree at least 2 is a path (we call it a backbone);

Theorem 5 (BL, Ochem, Pinlou, 2018)

(*i*)
$$\operatorname{RT}(2, CP) = 3;$$

(*ii*) $\operatorname{RT}(3, CP) = 2;$
(*iii*) $\operatorname{RT}(k, CP) = \frac{3}{2}, \text{ for } k \ge 4.$

RT of Subcubic Caterpillars - \mathcal{CP}_3

 Bounding the degree of caterpillars to 3 changes the behavior when the alphabet sizes grow;

RT of Subcubic Caterpillars - \mathcal{CP}_3

 Bounding the degree of caterpillars to 3 changes the behavior when the alphabet sizes grow;

Theorem 6 (BL, Ochem, Pinlou, 2018)

(i)
$$\operatorname{RT}(2, \mathcal{CP}_3) = 3;$$

(ii) $\operatorname{RT}(3, \mathcal{CP}_3) = 2;$
(iii) $\operatorname{RT}(4, \mathcal{CP}_3) = \frac{3}{2};$
(iv) $\operatorname{RT}(5, \mathcal{CP}_3) = \frac{4}{3};$
(v) $\operatorname{RT}(k, \mathcal{CP}_3) = 1 + \frac{1}{\lceil \frac{k}{2} \rceil}, \text{ for } k \ge 6.$

First, prove $\operatorname{RT}(2, \mathcal{CP}) \leq 3$:

■ There is a 2⁺-free coloring of backbone

First, prove $\operatorname{RT}(2, \mathcal{CP}) \leq 3$:

- There is a 2⁺-free coloring of backbone
- Color every leaf with a color distinct from its neighbor

Second, prove $\operatorname{RT}(2, \mathcal{CP}_3) \ge 3$: Two colors $x, y \in \{0, 1\}$

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- Suppose there is factor xyxy on backbone

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- Suppose there is factor xyxy on backbone



- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- Suppose there is factor xyxy on backbone



- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- Suppose there is factor xyxy on backbone



- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- Suppose there is factor xyxy on backbone



- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- Suppose there is factor xyxy on backbone



- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- Suppose there is factor xyx on backbone

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- Suppose there is factor xyx on backbone



- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- Suppose there is factor xyx on backbone



$\overline{\text{Proof of } \text{RT}(2, \mathcal{CP})} = \text{RT}(2, \mathcal{CP}_3) = 3$

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- Suppose there is factor xyx on backbone



- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- Suppose there is factor xyx on backbone


Proof of $\operatorname{RT}(2, \mathcal{CP}) = \operatorname{RT}(2, \mathcal{CP}_3) = 3$

Second, prove $\operatorname{RT}(2, \mathcal{CP}_3) \geq 3$:

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- Suppose there is factor xyx on backbone



Proof of $\operatorname{RT}(2, \mathcal{CP}) = \operatorname{RT}(2, \mathcal{CP}_3) = 3$

Second, prove $\operatorname{RT}(2, \mathcal{CP}_3) \geq 3$:

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- \Rightarrow there is no factor *xyx* on backbone

Proof of $\operatorname{RT}(2, \mathcal{CP}) = \operatorname{RT}(2, \mathcal{CP}_3) = 3$

Second, prove $\operatorname{RT}(2, \mathcal{CP}_3) \geq 3$:

- Two colors $x, y \in \{0, 1\}$
- There is no factor xxx
- \Rightarrow there is no factor *xyxy* on backbone
- \Rightarrow there is no factor *xyx* on backbone
- So, the coloring of backbone is comprised of consecutive factors xxyy, a contradiction

RT of Subcubic Trees - \mathcal{T}_3

 Similarly, bounding the degree of trees introduces additional hard problems

RT of Subcubic Trees - \mathcal{T}_3

 Similarly, bounding the degree of trees introduces additional hard problems

Theorem 7 (BL, Ochem, Pinlou, 2018)

(*i*) RT(4,
$$\mathcal{T}_3$$
) = $\frac{3}{2}$;
(*ii*) RT(5, \mathcal{T}_3) = $\frac{3}{2}$;
(*iii*) RT(k, \mathcal{T}_3) = $1 + \frac{1}{2\log k} + o\left(\frac{1}{\log k}\right)$, for $k \ge 6$.

RT of Subcubic Trees - \mathcal{T}_3

 Similarly, bounding the degree of trees introduces additional hard problems

Theorem 7 (BL, Ochem, Pinlou, 2018)

(i)
$$\operatorname{RT}(4, \mathcal{T}_3) = \frac{3}{2};$$

(ii) $\operatorname{RT}(5, \mathcal{T}_3) = \frac{3}{2};$
(iii) $\operatorname{RT}(k, \mathcal{T}_3) = 1 + \frac{1}{2\log k} + o\left(\frac{1}{\log k}\right), \text{ for } k \ge 6.$

• Open: k = 2 (known: $3 \le \operatorname{RT}(2, \mathcal{T}_3) \le \frac{7}{2}$) • Open: k = 3 (known: $2 \le \operatorname{RT}(3, \mathcal{T}_3) \le 3$)

Summary

	$ \mathbb{A} = 2$	$ \mathbb{A} = 3$	$ \mathbb{A} = 4$	$ \mathbb{A} = 5$	$ \mathbb{A} = k, \ k \ge 6$
\mathcal{P}	2	$\frac{7}{4}$	$\frac{7}{5}$	<u>5</u> 4	$\frac{k}{k-1}$
\mathcal{C}	<u>5</u> 2	2	$\frac{3}{2}$	$\frac{4}{3}$	$1+rac{1}{\lceilrac{k}{2} ceil}$
S	$\frac{7}{3}$	$\frac{7}{4}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$
\mathcal{CP}_3	3	2	$\frac{3}{2}$	$\frac{4}{3}$	$1+rac{1}{\lceilrac{k}{2} ceil}$
\mathcal{T}_3	?	?	$\frac{3}{2}$	$\frac{3}{2}$	$1 + \frac{1}{2\log k} + o\left(\frac{1}{\log k}\right)$
\mathcal{CP}	3	2	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$
$\overline{\mathcal{T}}$	$\frac{7}{2}$	3	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$

[1] ABERKANE, A., AND CURRIE, J. D.

There exist binary circular $5/2^+$ power free words of every length. *Electron. J. Combin. 11* (2004), #R10.

[2] Alon, N., Grytczuk, J., Haluszczak, M., and Riordan, O.

Non-repetitive colorings of graphs.

Random Struct. Algor. 21 (2002), 336-346.

[3] CARPI, A.

On Dejean's conjecture over large alphabets.

Theoret. Comput. Sci. 385, 1-3 (2007), 137-151.

Π

[4] CURRIE, J. D.

Non-repetitive Walks in Graphs and Digraphs. PhD thesis, University of Calgary, Alberta, Canada, 1987.

[5] CURRIE, J. D.

Which graphs allow infinite nonrepetitive walks? *Discrete Math. 87* (1991), 249–260.

[6] CURRIE, J. D.

There are ternary circular square-free words of length n for $n \ge 18$. Electron. J. Combin. 9 (2002), 1–7.

III

[7] CURRIE, J. D., MOL, L., AND RAMPERSAD, N.

Circular repetition thresholds on some small alphabets: Last cases of gorbunova's conjecture.

Electron. J. Combin. 26, 2 (2019), #P2.31.

[8] CURRIE, J. D., AND RAMPERSAD, N.

Dejean's conjecture holds for $n \ge 27$.

RAIRO - Theoretical Informatics and Applications 43, 4 (2009), 775–778.

[9] CURRIE, J. D., AND RAMPERSAD, N.

Dejean's conjecture holds for $n \ge 30$.

Theoret. Comput. Sci. 410, 30-32 (2009), 2885-2888.

IV

[10] CURRIE, J. D., AND RAMPERSAD, N.

A proof of Dejean's conjecture.

Math. Comp. 80 (2011), 1063-1070.

[11] DEJEAN, F.

Sur un théorème de Thue.

J. Combin. Theory Ser. A 13 (1972), 90-99.

[12] GORBUNOVA, I. A.

Repetition threshold for circular words.

Electron. J. Combin. 19, 4 (2012), P11.

[13] MOHAMMAD-NOORI, M., AND CURRIE, J. D.
 Dejean's conjecture and Sturmian words.
 Europ. J. Combin. 28, 3 (2007), 876–890.

[14] MOULIN OLLAGNIER, J.

Proof of Dejean's conjecture for alphabets with 5, 6, 7, 8, 9, 10 and 11 letters.

Theoret. Comput. Sci. 95, 2 (1992), 187-205.

[15] OCHEM, P., AND VASLET, E.

Repetition thresholds for subdivided graphs and trees.

RAIRO - Theoretical Informatics and Applications 46, 1 (2012), 123–130.

[16] PANSIOT, J.-J.

A propos d'une conjecture de F. Dejean sur les répétitions dans les mots.

```
Discrete Appl. Math. 7, 3 (1984), 297-311.
```

VI

[17] RAO, M.

Last cases of Dejean's conjecture.

Theoret. Comput. Sci. 412, 27 (2011), 3010-3018.

[18] THUE, A.

Über unendliche Zeichenreichen.

Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiana 7 (1906), 1-22.

Thank you!