On periodicities in strings

Frantisek (Franya) Franek

Department of Computing and Software
McMaster University, Hamilton, Ontario, Canada

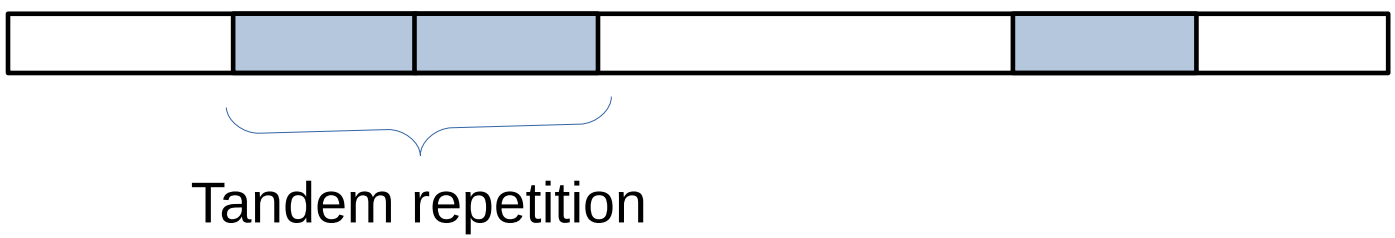
Invited talk at
Prague Stringology Conference 2025
Aug. 25-27, 2025

Outline

McMaster University

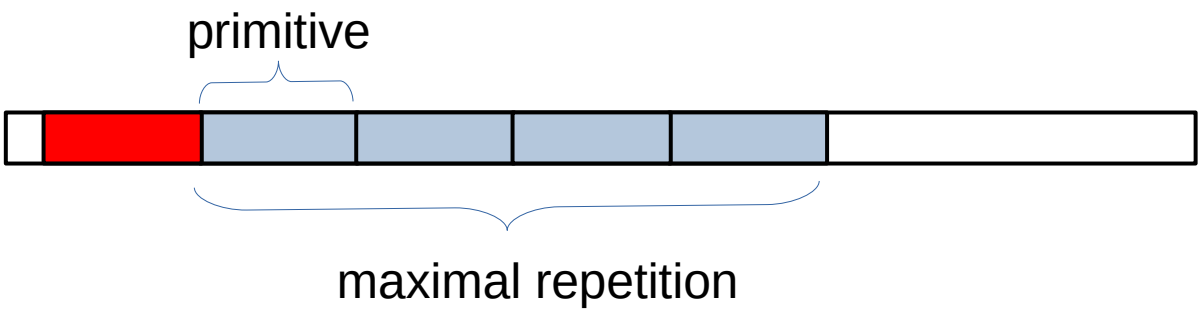
Introduction

- Tandem repetitions have been of interest to researchers in stringology from the beginning
- see for instance Fine & Wilf, *Uniqueness theorems for periodic functions*, 1965, or Berstel & Perrin, *The origins of combinatorics on words*, 2007



McMaster University

- The pioneering work of Crochemore in 1981 showed that the optimal bound for the number of (maximal) repetitions in a string of length n is of $O(n \log(n))$ complexity and attained by Fibonacci strings
- Crochemore, *An optimal algorithm for computing the repetitions in a word*, 1981
- *maximal repetition* in contemporary terminology would be called *primitively rooted leftmost maximal repetition*



McMaster University

- Followed closely by the seminal work by Apostolico and Preparata, in 1983, and Main in 1989 giving an $O(n \log(n))$ algorithm to detect all primitively rooted leftmost maximal repetitions.
- Apostolico & Preparata, *Optimal off-line detection of repetitions in a string*, 1983, Main, *Detecting leftmost maximal periodicities*,1989
- The $\log(n)$ factor in Main's algorithm comes from undetermined size of the alphabet. For a constant alphabet, it is linear.

McMaster University

This started intensive research in three areas:

- Which extension of the leftmost maximal primitively rooted repetition concept has a chance to be “linear” (i.e. linear in numbers and detectable by a linear algorithm) – this culminated in the concept of runs, and the maximum number of runs conjecture (in short runs conjecture).
- How many “squares” and “distinct squares” can a string contain – this culminated in the maximum number of distinct squares conjecture (in short distinct squares conjecture).

McMaster University

- Determining the combinatorial properties of many squares clustering – i.e. squares with the starting points in near proximity. Mostly with the aim to help resolve either the runs conjecture or the distinct squares conjecture.

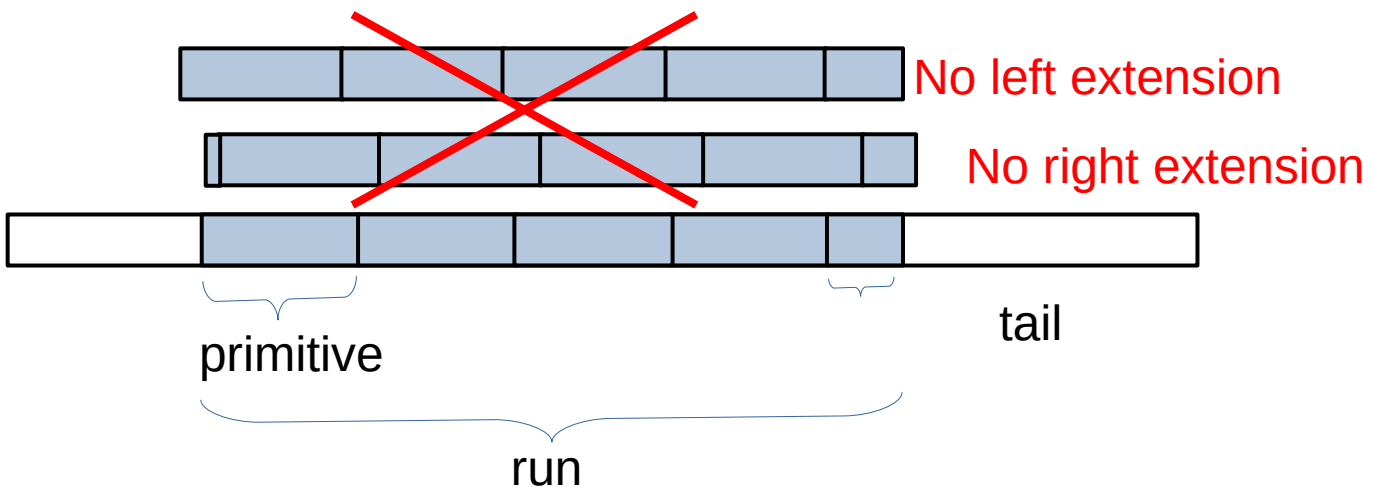
  Despite many interesting and important results obtained, only the early results of Crochemore & Ritter impacted indirectly the runs conjecture and directly the distinct squares conjecture.

  Crochemore & Rytter, *Squares, cubes, and time-space efficient strings searching*,1995

McMaster University

Runs

- Generalization of leftmost maximal repetitions (bundling them up).
- The term “runs” was coined by Smyth & Iliopoulos
- Iliopoulos & Moore & Smyth, *A linear algorithm for computing all the squares of a Fibonacci string*, 1996



Formally: **run** in a string $\boldsymbol{x} = \boldsymbol{x}[1 \,..\, n]$ is a four-tuple of integers $(start, period, exponent, tail)$ where

- $1 \leq s < n$, and
- $1 \leq p < n$, and
- $2 \leq e \leq n$, and
- $0 \leq t < p$, and
- $\boldsymbol{x}[s \,..\, ep{-}1{+}t]$ has minimal period p,
- either $s = 1$ or $\boldsymbol{x}[s{-}1 \,..\, ep{-}1{+}t]$ does not have period p (no left extension), and
- either $ep{-}1{+}t = n$ or $\boldsymbol{x}[s \,..\, ep{+}t]$ does not have period p (no right extension)

McMaster University

- Problem: bounding the number of occurrences of all runs.
- Runs may contain several squares bundled up, so bounding the maximum number of occurrences of squares is a different problem.
- In 1998, Fraenkel & Simpson showed that the number of occurrences of squares is bounded by $n \log_\Phi(n) \approx 1.441 n \log_2(n)$ (Φ *denotes the golden ratio*)
- Fraenkel & Simpson, *How many squares can a string contain?*, 1998

McMaster University

- The bound was improved in 2020 by Bannai et al. to $n \log_2(n)$.
- Bannai & Mieno & Nakashima, *Lyndon Words, the Three Squares Lemma, and Primitive Squares*, 2020
- The first significant jab at the problem of the maximum number of runs in a string comes in 1999 by Kolpakov & Kucherov: linear time in the length of the string.
- Kolpakov & Kucherov, *Finding maximal repetitions in a word in linear time*, 1999
- They also formalized the runs conjecture: the number of runs in a string is bounded by the length of the string.

McMaster University

- Kolpakov & Kucherov results spurred on a fury of research.
- In 2012, Deza & Franek with their grad student Andrew Baker introduced and investigated a d-step conjecture for runs: the number of runs in a string of length n with d distinct letters is bounded by $n - d$.
- Baker & Deza & Franek, *A parameterized formulation for the maximum number of runs problem*, *On the structure of run-maximal strings*

McMaster University

- This is not just an “insignificant lowering of the bound”, but something fundamental – the role of the size of the alphabet, not just the length of the string.
- The $(d, n-d)$ table and investigation of the relationship among the “neighbouring” entries of the table gave several insights how other entries can be computed in a linear-programming-like fashion.
- This approach was inspired by d-step approach for diameter-maximal polytopes connected to 1957 Hirsh conjecture.

McMaster University

- $\Delta(d, n) \leq n - d$, where $\Delta(d, n)$ denote the maximum possible diameter over all (d, n)-polytopes, i.e. polytopes of dimension d with n facets.
- The values of $\Delta(d, n)$ presented in $(d, n - d)$ table.
- In 1967 Klee & Walkup showed the equivalency between the Hirsch conjecture and the d-step conjecture stating that $\Delta(d, 2d) \leq d$ for all $d \geq 2$.
- The Hirsch conjecture was disproved by Santos in 2012 by exhibiting a violation on the main diagonal with $(d, n) = (43, 86)$.
- A continuous analogue of the Hirsch conjecture for the curvature of polytopes was proposed by Deza et al. in 2009 considering the simplex and central-path following primal-dual interior point methods.

McMaster University

- $\rho_d(n)$ denotes the maximum number of runs over all strings of length n with d distinct symbols.
- runs conjecture: $(\forall n \geq 2)(\forall 1 \leq d \leq n)(\rho_d(n) \leq n)$
- d-step conjecture for runs:
  $(\forall n \geq 2)(\forall 1 \leq d \leq n)(\rho_d(n) \leq n - d)$
- More interesting than $\rho_d(n) \leq n - d$ are the relationships in the $(d, n - d)$ table.

McMaster University

The $(d, n-d)$ table for $\rho_d(n)$.

	$\cdots$	$n-d$	$\cdots$
$\cdot$	$\cdots$	$\cdot$	$\cdots$
$\cdot$	$\cdots$	$\cdot$	$\cdots$
d	$\cdots$	$\rho_d(n)$	$\cdots$
$\cdot$	$\cdots$	$\cdot$	$\cdots$
$\cdot$	$\cdots$	$\cdot$	$\cdots$

d \ $n-d$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
2	**2**	*2*	*3*	*4*	*5*	***5***	***6***	***7***	***8***	***8***	***10***	10	11	12	13	14	15
3	2	**3**	*3*	*4*	*5*	*6*	***6***	***7***	***8***	***9***	***10***	***11***	11	12	13	14	15
4	2	3	**4**	*4*	*5*	*6*	*7*	***7***	***8***	***9***	***10***	***11***	***12***	12	13	14	15
5	2	3	4	**5**	*5*	*6*	*7*	*8*	***8***	***9***	***10***	***11***	***12***	***13***	13	14	15
6	2	3	4	5	**6**	*6*	*7*	*8*	*9*	***9***	***10***	***11***	***12***	***13***	***14***	14	15
7	2	3	4	5	6	**7**	*7*	*8*	*9*	*10*	***10***	***11***	***12***	***13***	***14***	***15***	15
8	2	3	4	5	6	7	**8**	*8*	*9*	*10*	*11*	***11***	***12***	***13***	***14***	***15***	***16***
9	2	3	4	5	6	7	8	**9**	*9*	*10*	*11*	*12*	***12***	***13***	***14***	***15***	***16***
10	2	3	4	5	6	7	8	9	**10**	*10*	*11*	*12*	*13*	***13***	***14***	***15***	***16***
11	2	3	4	5	6	7	8	9	10	**11**	*11*	*12*	*13*	*14*	***14***	***15***	***16***
12	2	3	4	5	6	7	8	9	10	11	**12**	*12*	*13*	*14*	*15*	***15***	***16***
13	2	3	4	5	6	7	8	9	10	11	12	**13**	*13*	*14*	*15*	*16*	***16***
14	2	3	4	5	6	7	8	9	10	11	12	13	**14**	*14*	*15*	*16*	*17*
15	2	3	4	5	6	7	8	9	10	11	12	13	14	**15**	*15*	*16*	*17*
16	2	3	4	5	6	7	8	9	10	11	12	13	14	15	**16**	*16*	*17*
17	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	**17**	*17*
18	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	**18**

McMaster University

- $\rho_d(n) \leq \rho_{d+1}(n+1)$ for $n \geq d \geq 2$
- $\rho_d(n) \leq \rho_d(n+1)$ for $n \geq d \geq 2$
- $\rho_d(n) < \rho_{d+1}(n+2)$ for $n \geq d \geq 2$
- $\rho_d(n) = \rho_{d+1}(n+1)$ for $2d \geq n \geq d \geq 2$
- $\rho_d(n) \geq n - d$, $\rho_d(2d+1) \geq d$ and $\rho_d(2d+2) \geq d+1$ for $2d \geq n \geq d \geq 2$
- $\rho_{d-1}(2d-1) = \rho_{d-2}(2d-2) = \rho_{d-3}(2d-3)$ and $0 \leq \rho_d(2d) - \rho_{d-1}(2d-1) \leq 1$ for $d \geq 5$.

McMaster University

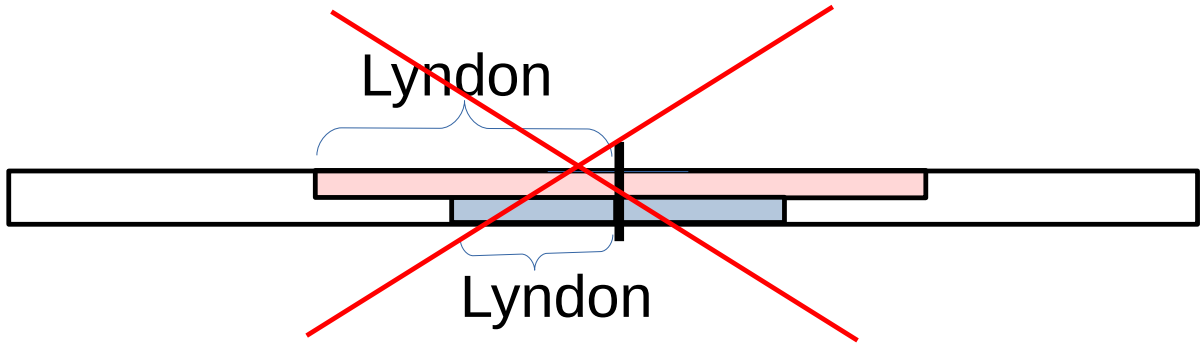
- $\rho_d(n) \leq n - d \iff \rho_d(2d) \leq d$
- $\rho_d(2d) = \rho_d(2d+1) \implies$ the string $a_1a_1a_2a_2 \cdots a_da_d$ is, up to a permutation of the alphabet, the unique run-maximal string of length $2d$ with d distinct symbols
- $\rho_d(2d+1) = \rho_d(2d+2) = \rho_d(2d+3)$.
- $n - d$ is an optimal bound (in terms of n and d) as strings of length n with d symbols having $n - d$ runs are known: string $a_1a_1a_2a_2 \cdots a_da_d$ has length $n = 2d$, has d distinct symbols, and has d runs.

McMaster University

- These relationships, together with R-cover, allowed computing binary run-maximal strings of length < 60 in hours verifying computational results by Kolpakov & Kucherov and extend it significantly to lengths < 73, and to compute values for higher alphabets that would be inaccessible by direct (brute force) approach.
- see `https://advol.cas.mcmaster.ca/bakerar2/research/runmax/index.html`
- Runs conjecture was proven by Bannai & I & Inenaga & Nakashima & Takeda & Tsuruta in 2015 mapping runs to their L-roots
- Bannai & I & Inenaga & Nakashima & Takeda & Tsuruta, *"Runs" Theorem*, 2015

McMaster University

- Crochemore (& Rytter ?) knew that if two squares are Lyndon (have Lyndon roots), they cannot share the midpoint:



- Thus each cube can be assigned the midpoint of the Lyndon square that is guaranteed to exist in it, as a consequence there are at most n cubes in a string of length n.

- There always is a single Lyndon root in a square. But the problem is that two Lyndon roots can start at the same point.
- The essence of Bannai et al. idea was the realization that they cannot be both maximal Lyndon roots, and they ingeniously came with the way to make them maximal Lyndon with respect to the given order or its inverse of the alphabet of the string.
- Thus every ran has one or more maximal Lyndon roots with respect to the given order or its inverse – what Bannai et al. call L-roots, so the maximum number of runs is bounded be the maximum number of L-roots, hence by the length of the string.

McMaster University

- In the same year, 2015, Deza and myself realized that the number of L-roots is bounded not by the length of the string, but by the d-step bound, and hence we proved the d-step conjecture for runs with all the relationships following from it.
- Deza & Franek, *Bannai et al. method proves the d-step conjecture for strings*
- Note that there is no wiggle room for $(d, 2d)$, while for binary strings the best upper bound is $\frac{22}{23}n < 0.957n$, due to Fischer & Holub & I & Lewenstein
- Fischer & Holub & I & Lewenstein,*Beyond the Runs theorem*, 2015

McMaster University

Distinct squares

- In maximum number of distinct squares problem, we count not the occurrences of squares, but their types.
- Thus *aabaab* has two distinct squares, *aa* and *aabaab*, though it has three occurrences of squares – *aa* is only counted once.
- This is a kind of problem not well suited for structural analysis, either one must decide which occurrence of the multiple occurrences of a square to count, or one must mapped all the squares to entities in a way that the occurrences of the same type of a square are mapped onto the same entity, and then count the entities instead of the squares.

- The problem was brought into a sharp focus of interest by the seminal work of Fraenkel & Simpson in 1997.
- Fraenkel & Simpson, *How Many Squares Can a String Contain?*, 1998
- The decided to count the **right-most** occurrences of squares. Employing a result of Crochemore & Rytter (later became known as Three Square Lemma) they showed that at most two right-most squares can start at the same position.

McMaster University

- They showed if three squares start at the same position, the smallest of the three squares would not be right-most.
- Thus concluded that the upper bound must be smaller than twice the length.
- They also introduced the distinct squares conjecture – the number of distinct squares is bounded by the length of the string and showed that it is asymptotically optimal.
- in 2005 Ilie simplified their proof eliminating need for a direct application of the Three Square Lemma.
- Ilie, *A simple proof that a word of length n has at most 2n distinct squares*, 2005

McMaster University

- The combinatorial/structural investigation of double-squares was initiated by Lam, and followed by Deza & Franek & Thierry, improving the bound to $\frac{11}{6}n$.
- Deza & Franek & Thierry, *How many double squares can a string contain?*, 2015

McMaster University

- $\sigma_d(n)$ denotes the maximum number of distinct squares over all strings of length n with d distinct symbols.
- Deza & Franek & Jaing, *A d-step approach for distinct squares in strings*, 2011, *A Computational Framework for Determining Square-maximal Strings*, 2012
- distinct square conjecture:
  $(\forall n \geq 2)(\forall 1 \leq d \leq n)(\sigma_d(n) \leq n)$
- *d*-step conjecture for distinct squares:
  $(\forall n \geq 2)(\forall 1 \leq d \leq n)(\sigma_d(n) \leq n - d)$

McMaster University

$(d, n-d)$ table

		$n-d$										
		1	2	3	4	5	6	7	8	9	10	11
	1	1	1	1	1	1	1	1	1	1	1	$\sigma_1(12)$
	2	1	2	2	3	3	4	5	6	7	7	$\sigma_2(13)$
	3	1	2	3	3	4	4	5	6	7	8	$\sigma_3(14)$
	4	1	2	3	4	4	5	5	6	7	8	$\sigma_4(15)$
	5	1	2	3	4	5	5	6	6	7	8	$\sigma_5(16)$
d	6	1	2	3	4	5	6	6	7	7	8	$\sigma_6(17)$
	7	1	2	3	4	5	6	7	7	8	8	$\sigma_7(18)$
	8	1	2	3	4	5	6	7	8	8	9	$\sigma_8(19)$
	9	1	2	3	4	5	6	7	8	9	9	$\sigma_9(20)$
	10	1	2	3	4	5	6	7	8	9	10	$\sigma_{10}(21)$
	11	$\sigma_{11}(12)$	$\sigma_{11}(13)$	.	.	.	.	.	.	.	.	.

The main diagonal, the second diagonal

For any $2 \leq d \leq n$:

(a) $\sigma_d(n) \leq \sigma_d(n+1)$
the values are non-decreasing when moving left-to-right along a row

(b) $\sigma_d(n) \leq \sigma_{d+1}(n+1)$
the values are non-decreasing when moving top-to-bottom along a column

(c) $\sigma_d(n) < \sigma_{d+1}(n+2)$
the values are strictly increasing when moving left-to-right and top-to-bottom along descending diagonals

McMaster University

(d) $\sigma_d(2d) = \sigma_d(n) = \sigma_{d+1}(n+1)$ for $n \leq 2d$
the values under and on the main diagonal along a column are constant

(e) $\sigma_d(n) \geq n-d$ for $n \leq 2d$
the values under and on the main diagonal are at least as big as conjectured: $\sigma_d(2d+1) \geq d$ *and* $\sigma_d(2d+2) \geq d+1$

(f) $\sigma_d(2d) - \sigma_{d-1}(2d-1) \leq 1$
the difference between the value on the main diagonal and the value immediately above it is no more than 1

McMaster University

Theorem (The main diagonal dominates)

$\sigma_d(n) \leq n-d$ *holds true for all* $2 \leq d \leq n$ *iff* $\sigma_d(2d) \leq d$ *for every* $d \geq 2$.

Theorem (If the main diagonal and the second one are "close")

$\sigma_d(n) \leq n-d$ *holds true for all* $2 \leq d \leq n$ *iff* $\sigma_d(2d+1)-\sigma_d(2d) \leq 1$ *for every* $d \geq 2$.

McMaster University

Theorem (If second diagonal bounded, a stronger upper bound)

If $\sigma_d(2d+1) \leq d$ *for every* $d \geq 2$, *then* $\sigma_d(n) \leq n-d-1$ *for* $n > 2d \geq 4$ *and* $\sigma_d(n) = n-d$ *for* $n \leq 2d$.

Theorem (If the main diagonal and the second one are the same, a stronger upper bound)

If $\sigma_d(2d) = \sigma_d(2d+1)$ *for every* $d \geq 2$, *then* $\sigma_d(n) \leq n-d-1$ *for* $n > 2d \geq 4$ *and* $\sigma_d(n) = n-d$ *for* $n \leq 2d$.

McMaster University

- Thus, to prove the conjecture in general, it is sufficient to prove it for the main diagonal.

- The sets of square-maximal strings of length n and d distinct symbols are irregular and unstructured, while the square-maximal strings on the main diagonal are regular and structured -- see Mei Jiang's website `https://advol.cas.mcmaster.ca/jiangm5/research/square.html`

McMaster University

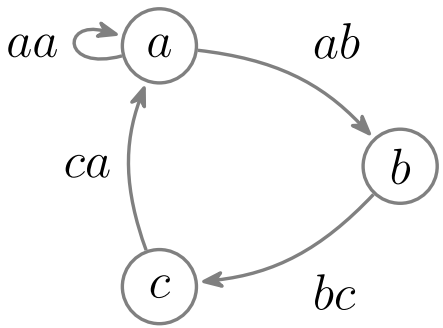
- The *d*-step conj. was recently proven by Brlek & Li
  - Brlek & Li, *On the number of squares in a finite word*, arXiv (2022)
  - Brlek & Li: *On the number of squares in a finite word*. Comb. Theory 5(1) (2025)
- Their method, unlike the proof of *d*-step conjecture for runs, clearly illustrates and illuminates the role of the size of the alphabet.
- For the distinct square problem, either you have to decide which occurrences of squares to count (Fraenkel & Simpson, Lam, Deza & Franek & Thierry), or map all the occurrences of squares on some entities so that the occurrences of the squares of the same type are mapped to the same entity, and count the entities instead (Brlek & Li).

This is a joint work with grad student Holly Kopponen

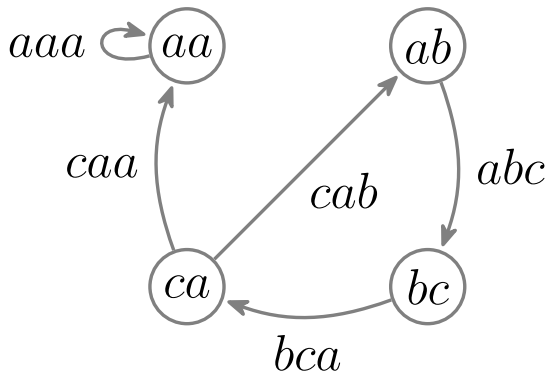
- Berge theorem: the cyclomatic number $|A| - |V| + 1$ of a weakly connected directed graph $G = (V, A)$ equals the number of independent directed cycles in G (see any book on graph theory by Berge)
- Map all occurrences of squares in a string $\boldsymbol{x}$ to so-called small directed cycles in the Rauzy graph of $\boldsymbol{x}$.
- Prove that the cycles in the set of small directed cycles in the Rauzy graph of $\boldsymbol{x}$ are independent.
- Apply the cyclomatic number theorem of Berge to the individual components of the Rauzy graph of $\boldsymbol{x}[1 \,..\, n]$ to obtain the upper bound of $n - d$ where d is the number of distinct symbols of $\boldsymbol{x}$.

McMaster University

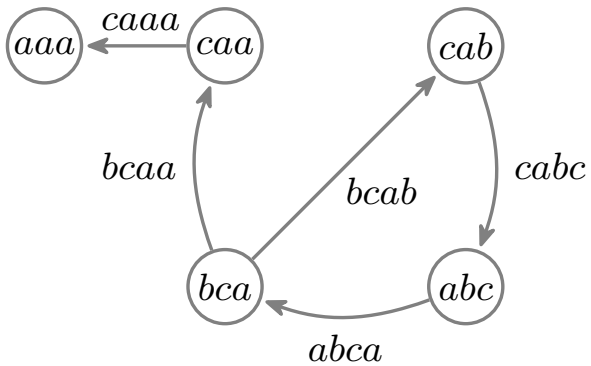
Rauzy graph of $\boldsymbol{x} = abcabcaaa$, components $R_1(\boldsymbol{x})$ .. $R_5(\boldsymbol{x})$:



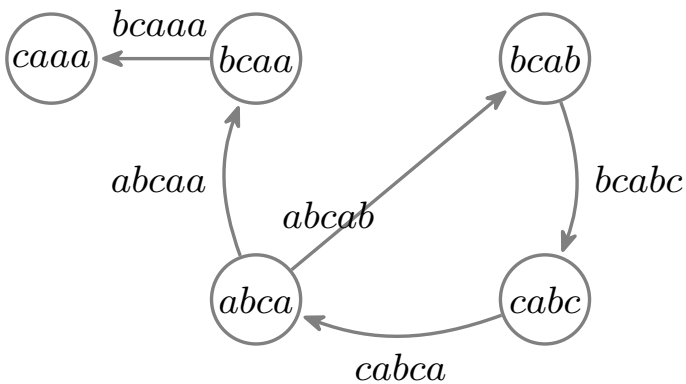
$\boldsymbol{R}_1(\boldsymbol{u})$



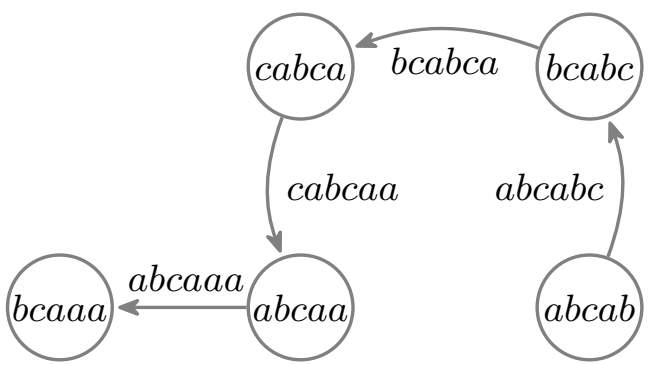
$\boldsymbol{R}_2(\boldsymbol{u})$



$\boldsymbol{R}_3(\boldsymbol{u})$

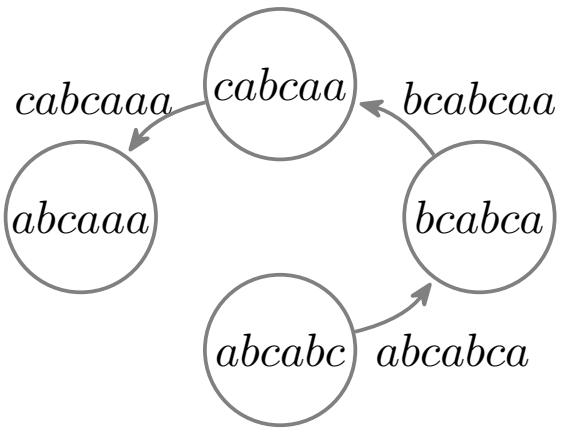


$\boldsymbol{R}_4(\boldsymbol{u})$

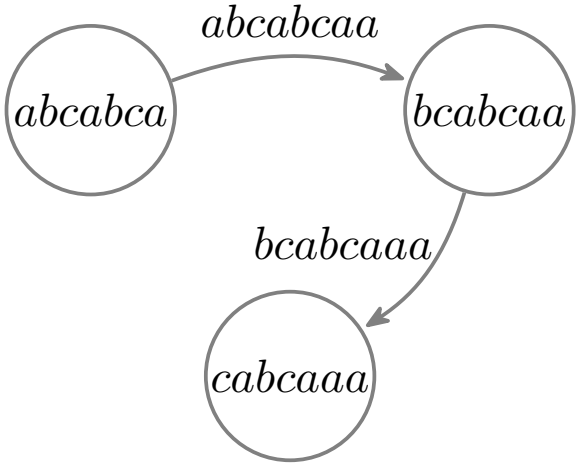


$\boldsymbol{R}_5(\boldsymbol{u})$

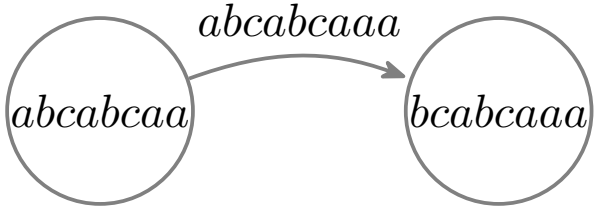
Rauzy graph of $\boldsymbol{x} = abcabcaaa$, components $R_6(\boldsymbol{x})$ .. $R_9(\boldsymbol{x})$:



$\boldsymbol{R_6(u)}$



$\boldsymbol{R_7(u)}$



$\boldsymbol{R_8(u)}$

$\boldsymbol{R_9(u)}$

- Consider as square *abc.abc* in $\boldsymbol{x}$, it corresponds to a (canonical) cycle of length 3 in $R_3(\boldsymbol{x})$:

$$abc \xrightarrow{abca} bca \xrightarrow{bcab} cab \xrightarrow{cabc} abc$$

  so we can map any occurrence of the square *abc.abc* in $\boldsymbol{x}$ to this cycle.

- But what if $\boldsymbol{x}$ contains somewhere else a square *cab.cab* – it would map canonically to the same cycle. But it can be mapped to a cycle of length 3 in the dimension one up

$$cabc \xrightarrow{cabca} abca \xrightarrow{abcab} bcab \xrightarrow{bcabc} cabc$$

  in $R_4(\boldsymbol{x})$ (*lifting via rank*)

McMaster University

- But what if the square is not primitively rooted? *ab.ab.ab.ab*

$$abab \xrightarrow{ababa} baba \xrightarrow{babab} abab$$

  a cycle of length 2 in $R_4(\boldsymbol{x})$.

- Note all the cycles are of length $\leq$ the dimension of the Rauzy component, i.e. so-called short cycles.

- To prove that the short cycles in a Rauzy component are independent is quite involved. That's what is in Brlek & Li papers.
- We opted for a more formal approach, defining a mapping and proving that the set of ϕ-cycles in a Rauzy component are independent, which is simpler.

Definition (Mapping $\phi_{\boldsymbol{x}}$)

Let $\boldsymbol{uu}$ be a square in a string $\boldsymbol{x}$ of rank r. Let $|\boldsymbol{u}| = k$ and let the minimal period of $\boldsymbol{u}$ be ℓ. Let $\widehat{\boldsymbol{u}}$ be the leading root. Then $\phi_{\boldsymbol{x}}(\boldsymbol{uu})$ is defined as the directed cycle in $R_{k+r-1}(\boldsymbol{x})$ of length ℓ whose set of vertices is $\{\mathfrak{rs}(\widehat{\boldsymbol{u}}\widehat{\boldsymbol{u}}\widehat{\boldsymbol{u}}, k+r-1, i) \mid 0 \leq i < \ell\}$ and whose set of arcs is $\{\mathfrak{rs}(\widehat{\boldsymbol{u}}\widehat{\boldsymbol{u}}\widehat{\boldsymbol{u}}, k+r, i) \mid 0 \leq i < \ell\}$.

- Hence the number of distinct squares with the primitive root of their roots of length r is equal to the number of ϕ-cycles in $R_r(\boldsymbol{x}) = ([\boldsymbol{x}]_r, [\boldsymbol{x}]_{r+1})$, majorized by $|[\boldsymbol{x}]_{r+1}| - |[\boldsymbol{x}]_r| + 1$ by Berge theorem.
- Thus, over all components $R_1(\boldsymbol{x}) \dots R_{n-1}(\boldsymbol{x})$):

$$(|[\boldsymbol{x}]_2| - |[\boldsymbol{x}]_1| + 1) + (|[\boldsymbol{x}]_3| - |[\boldsymbol{x}]_2| + 1) + \ldots + (|[\boldsymbol{x}]_n| - |[\boldsymbol{x}]_{n-1}| + 1) = \underbrace{1 + \ldots + 1}_{n-1 \text{ times}} + \underbrace{|[\boldsymbol{x}]_n|}_{=1} - \underbrace{|[\boldsymbol{x}]_1|}_{=d} = n - d.$$

A few remaining open problems

- The conjecture of Jonoska & Manea & Seki for binary strings is still unresolved: $\sigma_2(n) \leq \frac{2k-1}{2k+2}n$
  where k is the number of occurrences of the letter with the smaller frequency of occurrence.

- $k \leq \frac{n}{2}$, and so $\frac{2k-1}{2k+2}n < n-2$ and so it is stronger than d-step conjecture.

- Jonoska & Manea & Seki, *A Stronger Square Conjecture on Binary Words*, 2014

McMaster University

- The computational work done on the $(d, n-d)$ tables for runs and for distinct squares indicated $\sigma_d(n) \leq \rho_d(n)$.

- see `https://advol.cas.mcmaster.ca/bakerar2/research/runmax/index.html`

  see `https://advol.cas.mcmaster.ca/jiangm5/research/square.html`

McMaster University

- There is no relationship between $\boldsymbol{r}(\boldsymbol{x})$ and $\boldsymbol{s}(\boldsymbol{x})$: for instance $\boldsymbol{x} = abcabcabc$ has 3 distinct squares *abcabc*, *bcabca* and *cabcab* while it only has 1 run *abcabcabc*.
- $\boldsymbol{x} = ababxabab$ has 1 distinct square *abab* while it has 2 runs *abab* starting at position 1 and *abab* starting at position 6.
- Nevertheless all computed values of $\sigma_d(n)$ and $\rho_d(n)$ indicate that $\sigma_d(n) \leq \rho_d(n)$ which Deza and I conjecture to be true for all admissible values of n and d.

McMaster University

THANK YOU

McMaster University