Words of minimum rank in deterministic finite automata

Jarkko Kari [1] Andrew Ryzhikov [2] Anton Varonka [3]

[1] University of Turku

[2] LIGM, Université Paris-Est Marne-la-Vallée

[3] Belarusian State University

August 6th, 2019

Finite automata

Finite (complete deterministic) automaton:

- $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$
- Q and Σ are both finite and non-empty
- $\delta : Q \times \Sigma \to Q$ is a complete transition function
- no initial or final states
- words $w \in \Sigma^*$ act on Q

Finite automata

Finite (complete deterministic) automaton:

- $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$
- Q and Σ are both finite and non-empty
- $\delta : Q \times \Sigma \to Q$ is a complete transition function
- no initial or final states
- words $w \in \Sigma^*$ act on Q

Synchronizing automata

An automaton is called *synchronizing* if there exists a word $w \in \Sigma^*$ such that $\delta(q, w) = \delta(q', w)$ for any $q, q' \in Q$.
Any such word w is said to be *synchronizing* for the given automaton.

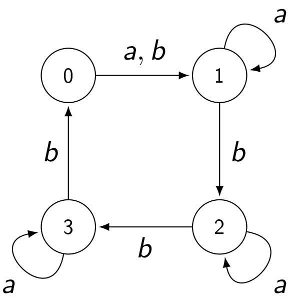


Figure: An automaton synchronized by ab^3ab^3a.

How synchronization takes place

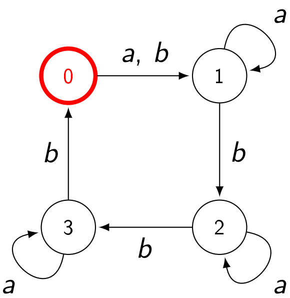


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

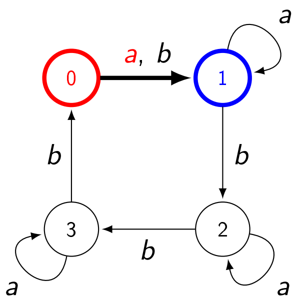


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

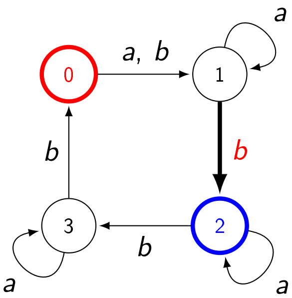


Figure: The automaton is synchronized by a word $abbbabbba$.

How synchronization takes place

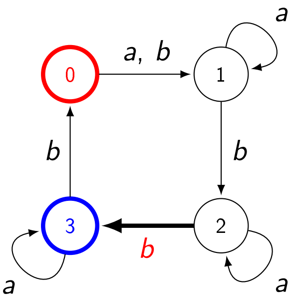


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

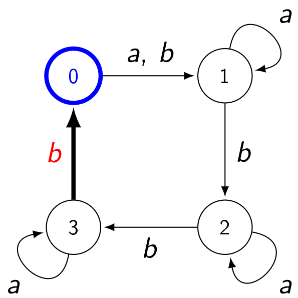


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

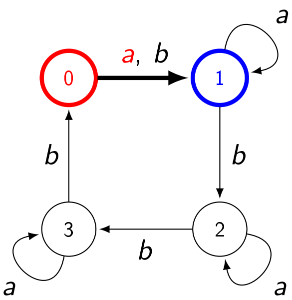


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

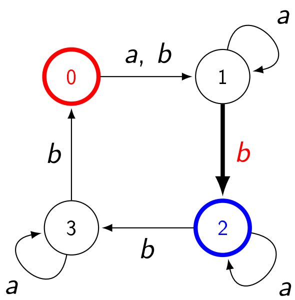


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

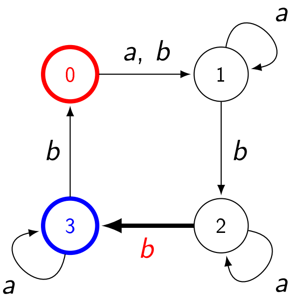


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

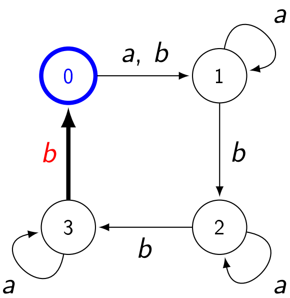


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

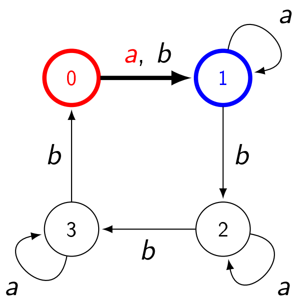


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

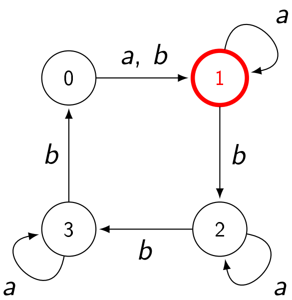


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

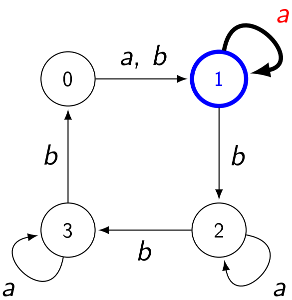


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

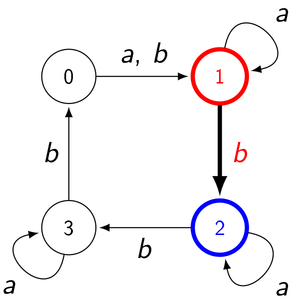


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

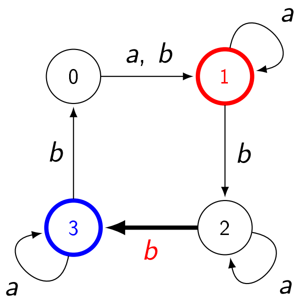


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

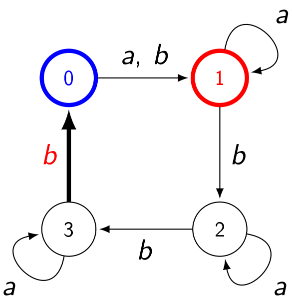


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

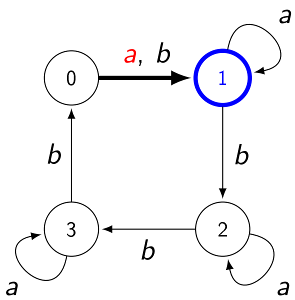


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

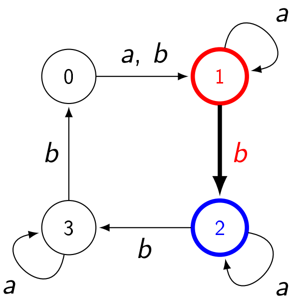


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

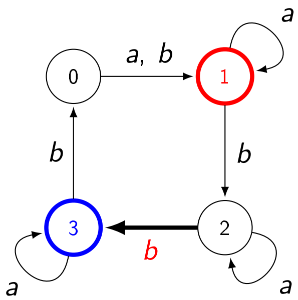


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

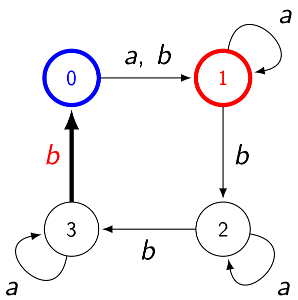


Figure: The automaton is synchronized by a word *abbbabbba*.

How synchronization takes place

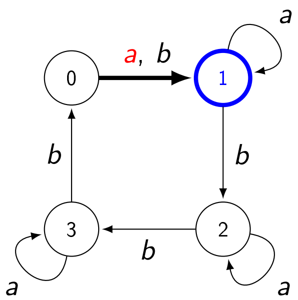


Figure: The automaton is synchronized by a word *abbbabbba*.

The Černý Conjecture

- minimum length of a synchronizing word for $\mathcal{A}$ is called the reset threshold and is denoted $\mathrm{rt}(\mathcal{A})$.

The Černý Conjecture, [2].

Every synchronizing automaton with n states can be synchronized by a word of length

$$\mathrm{rt}(\mathcal{A}) \leq (n-1)^2.$$

- this bound is achieved for automata of $\mathcal{C}_n$ series.

Have you ever seen the rank?

- a synchronizing word «merges» all states to one particular
- the *rank* of a word w is equal to the number of states active after applying it, i.e. $|\{\delta(Q, w)\}|$
- the *rank* of an automaton is the minimum rank among all words $w \in \Sigma^*$
- a synchronizing word (automaton) is a word (automaton) of rank 1
- rank of a strongly connected automaton is bounded from below by the *period* — g.c.d. of cycles lengths in an underlying digraph

Have you ever seen the rank?

- a synchronizing word «merges» all states to one particular
- the *rank* of a word w is equal to the number of states active after applying it, i.e. $|\{\delta(Q, w)\}|$
- the *rank* of an automaton is the minimum rank among all words $w \in \Sigma^*$
- a synchronizing word (automaton) is a word (automaton) of rank 1
- rank of a strongly connected automaton is bounded from below by the *period* — g.c.d. of cycles lengths in an underlying digraph

Have you ever seen the rank?

- a synchronizing word «merges» all states to one particular
- the *rank* of a word w is equal to the number of states active after applying it, i.e. $|\{\delta(Q, w)\}|$
- the *rank* of an automaton is the minimum rank among all words $w \in \Sigma^*$
- a synchronizing word (automaton) is a word (automaton) of rank 1
- rank of a strongly connected automaton is bounded from below by the *period* — g.c.d. of cycles lengths in an underlying digraph

Have you ever seen the rank?

- a synchronizing word «merges» all states to one particular
- the *rank* of a word w is equal to the number of states active after applying it, i.e. $|\{\delta(Q, w)\}|$
- the *rank* of an automaton is the minimum rank among all words $w \in \Sigma^*$
- a synchronizing word (automaton) is a word (automaton) of rank 1
- rank of a strongly connected automaton is bounded from below by the *period* — g.c.d. of cycles lengths in an underlying digraph

Have you ever seen the rank?

- a synchronizing word «merges» all states to one particular
- the *rank* of a word w is equal to the number of states active after applying it, i.e. $|\{\delta(Q, w)\}|$
- the *rank* of an automaton is the minimum rank among all words $w \in \Sigma^*$
- a synchronizing word (automaton) is a word (automaton) of rank 1
- rank of a strongly connected automaton is bounded from below by the *period* — g.c.d. of cycles lengths in an underlying digraph

Minimum rank threshold

$\mathrm{mrt}(\mathcal{A})$ — *minimum rank threshold* — is the length of a shortest word of minimum rank of an automaton $\mathcal{A}$.

The Rank Conjecture (Pin, Pribavkina)

Let $\mathcal{A}$ be an n-state automaton of rank r. Then

$$\mathrm{mrt}(\mathcal{A}) \leq (n-r)^2.$$

- Pin [3] provided an example of n-automaton of rank r with $\mathrm{mrt}(\mathcal{A}) = (n-r)^2$
- is in fact an automaton $\mathcal{C}_{n-r+1}$ with $(r-1)$ isolated states
- series of strongly connected automata with $\mathrm{mrt}(\cdot)$ close to that bound?

Minimum rank threshold

$\mathrm{mrt}(\mathcal{A})$ — *minimum rank threshold* — is the length of a shortest word of minimum rank of an automaton $\mathcal{A}$.

The Rank Conjecture (Pin, Pribavkina)

Let $\mathcal{A}$ be an n-state automaton of rank r. Then

$$\mathrm{mrt}(\mathcal{A}) \leq (n-r)^2.$$

- Pin [3] provided an example of n-automaton of rank r with $\mathrm{mrt}(\mathcal{A}) = (n-r)^2$
- is in fact an automaton $\mathcal{C}_{n-r+1}$ with $(r-1)$ isolated states
- series of strongly connected automata with $\mathrm{mrt}(\cdot)$ close to that bound?

Minimum rank threshold

$\mathrm{mrt}(\mathcal{A})$ — *minimum rank threshold* — is the length of a shortest word of minimum rank of an automaton $\mathcal{A}$.

The Rank Conjecture (Pin, Pribavkina)

Let $\mathcal{A}$ be an n-state automaton of rank r. Then

$$\mathrm{mrt}(\mathcal{A}) \leq (n-r)^2.$$

- Pin [3] provided an example of n-automaton of rank r with $\mathrm{mrt}(\mathcal{A}) = (n-r)^2$
- is in fact an automaton $\mathcal{C}_{n-r+1}$ with $(r-1)$ isolated states
- series of strongly connected automata with $\mathrm{mrt}(\cdot)$ close to that bound?

Lower bounds on mrt(·): ternary automata

Proposition 1.

For every n and every $r > 1$ such that r divides n, there exists a ternary strongly connected automaton with n states and rank r such that the length of its shortest word of minimum rank is $\frac{(n-r)^2}{r} + r - 1$.

- binarization techniques yield a lower bound of $\mathrm{mrt}(\mathcal{A}) \geq \frac{n^2}{3r} - \frac{7}{3}n + 5r$ for binary automata.

Lower bounds on mrt(·): ternary automata

Proposition 1.

For every n and every $r > 1$ such that r divides n, there exists a ternary strongly connected automaton with n states and rank r such that the length of its shortest word of minimum rank is $\frac{(n-r)^2}{r} + r - 1$.

- binarization techniques yield a lower bound of $\mathrm{mrt}(\mathcal{A}) \geq \frac{n^2}{3r} - \frac{7}{3}n + 5r$ for binary automata.

Lower bounds on mrt(·): binary automata

Proposition 2

For every t-state strongly connected synchronizing automaton $\mathcal{A}$ and for every r there exists a tr-state strongly connected automaton $\mathcal{A}'$ over the same alphabet, with period and rank equal to r, such that $\mathrm{mrt}(\mathcal{A}') = \mathrm{rt}(\mathcal{A}) \cdot r$.

Corollary 1

For every n and every r such that r divides n, there exists an n-state circular[1] binary automaton of period and rank r with minimum rank threshold $\frac{(n-r)^2}{r}$.

[1] with a letter acting as a cyclic permutation

Lower bounds on mrt(·): binary automata

Hybrid Rank-Road Coloring problem

Given a fixed digraph with n vertices and period p determine the shortest length of a word of rank p among all colorings of it.

Corollary 2

For every $n > 1$ and every r such that r divides n, there exists an n-vertex strongly connected digraph D of constant outdegree 2 such that all colorings of D are circular, have the same period and rank r, and for every coloring the length of a word of minimum rank is $\frac{(n-r)^2}{r} - n + 2r$.

Minimum rank threshold and index

Recall that the *index* of a strongly connected digraph with period r is the smallest k such that any pair of vertices are connected by a directed path of length k if and only if they are connected by a path of length $k + r$.

Proposition 3

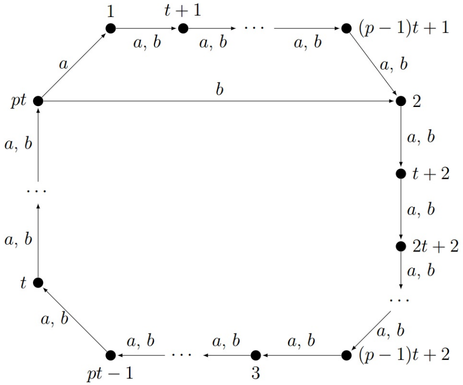
Let $\mathcal{A}$ be a strongly connected n-state automaton of rank r and period r. Then the following holds:

$$\mathrm{mrt}(\mathcal{A}) \geq k(\mathcal{A}) - n + r,$$

where $k(\mathcal{A})$ is the index of the underlying digraph of $\mathcal{A}$.

This generalizes the result of [5] for the case $r > 1$.

Modified Wielandt automaton



$$\mathrm{mrt}(\mathcal{A}_{n,r}) = \frac{(n-r)^2}{r} - n + 2r$$

Another lower bound on mrt(·)

Proposition 4

For every n and every $r > 1$ such that r divides n, there exists a binary n-state circular automaton $\mathcal{A}$ of rank r having $\mathrm{mrt}(\mathcal{A}) = \frac{(n-r)^2}{r} + r - 1$.

- same lower bound for binary automata as in the ternary case (Proposition 1)

Upper bounds on $\mathrm{mrt}(\cdot)$

For every n, let $f(n)$ denote the maximum of reset thresholds of n-state synchronizing automata.

Theorem 1

Let $\mathcal{A}$ be an automaton of rank r and period r. Then $\mathrm{mrt}(\mathcal{A}) \leq r^2 \cdot f(\frac{n}{r}) + (r-1)$.

- implies an unconditional bound $\mathrm{mrt}(\mathcal{A}) \leq \frac{n(n^2-r^2)}{6r} + r - 1$
- the Černý Conjecture would imply $\mathrm{mrt}(\mathcal{A}) \leq (n-r)^2 + r - 1$
- we conjecture the upper bound can be improved to $\frac{(n-r)^2}{r} + O(n)$

Upper bounds on mrt(·)

For every n, let $f(n)$ denote the maximum of reset thresholds of n-state synchronizing automata.

Theorem 1

Let $\mathcal{A}$ be an automaton of rank r and period r. Then $\mathrm{mrt}(\mathcal{A}) \leq r^2 \cdot f(\frac{n}{r}) + (r-1)$.

- implies an unconditional bound $\mathrm{mrt}(\mathcal{A}) \leq \frac{n(n^2-r^2)}{6r} + r - 1$
- the Černý Conjecture would imply $\mathrm{mrt}(\mathcal{A}) \leq (n-r)^2 + r - 1$
- we conjecture the upper bound can be improved to $\frac{(n-r)^2}{r} + O(n)$

Upper bounds on $\mathrm{mrt}(\cdot)$

For every n, let $f(n)$ denote the maximum of reset thresholds of n-state synchronizing automata.

Theorem 1

Let $\mathcal{A}$ be an automaton of rank r and period r. Then $\mathrm{mrt}(\mathcal{A}) \leq r^2 \cdot f(\frac{n}{r}) + (r-1)$.

- implies an unconditional bound $\mathrm{mrt}(\mathcal{A}) \leq \frac{n(n^2-r^2)}{6r} + r - 1$
- the Černý Conjecture would imply $\mathrm{mrt}(\mathcal{A}) \leq (n-r)^2 + r - 1$
- we conjecture the upper bound can be improved to $\frac{(n-r)^2}{r} + O(n)$

Eulerian automata

- an automaton is *Eulerian* if it is strongly connected and its underlying digraph is Eulerian
- Kari [4] showed that $\mathrm{rt}(\mathcal{A}) \leq (n-1)(n-2)+1$ for such automata
- we extend that result to the case of arbitrary minimum rank

Eulerian automata

- an automaton is *Eulerian* if it is strongly connected and its underlying digraph is Eulerian
- Kari [4] showed that $\mathrm{rt}(\mathcal{A}) \leq (n-1)(n-2)+1$ for such automata
- we extend that result to the case of arbitrary minimum rank

Eulerian automata

- an automaton is *Eulerian* if it is strongly connected and its underlying digraph is Eulerian
- Kari [4] showed that $\mathrm{rt}(\mathcal{A}) \leq (n-1)(n-2)+1$ for such automata
- we extend that result to the case of arbitrary minimum rank

The Rank Conjecture for Eulerian automata

Theorem 2.

Let $\mathcal{A}$ be an n-state Eulerian automaton of rank r. Then $\mathcal{A}$ has a word of rank r of length at most $(n-r-1)(n-r)+1$.

Proof sketch.

- Key idea inherited from [4]: subsets of states are considered as vectors in $\mathbb{R}^n$.
- Further, a linear weight function is defined as follows: $|(x_1,\ldots,x_n)| = x_1+\ldots+x_n$.
- Define $f_w^{-1}(q) = \{q' \mid \delta(q',w) = q\}$. This is furthermore extended to a linear mapping $f_w^{-1} : \mathbb{R}^n \to \mathbb{R}^n$.

The Rank Conjecture for Eulerian automata

Theorem 2.

Let $\mathcal{A}$ be an n-state Eulerian automaton of rank r. Then $\mathcal{A}$ has a word of rank r of length at most $(n-r-1)(n-r)+1$.

Proof sketch.

- Key idea inherited from [4]: subsets of states are considered as vectors in $\mathbb{R}^n$.
- Further, a linear weight function is defined as follows: $|(x_1, \ldots, x_n)| = x_1 + \ldots + x_n$.
- Define $f_w^{-1}(q) = \{q' \mid \delta(q', w) = q\}$. This is furthermore extended to a linear mapping $f_w^{-1} : \mathbb{R}^n \to \mathbb{R}^n$.

Proof sketch continued

- Some vectors are *non-extendable*, for such vector x there exists no word w with $|f_w^{-1}(x)| \neq |x|$.
- The set of non-extendable vectors is a linear subspace of $\mathbb{R}^n$ with dimension *at least* r. We denote it by Z_1 and its orthogonal complement by Z_0.
- Any vector can be decomposed into $x = x_0 + x_1$ with $x_i \in Z_i$. Then one can prove that for every extendable vector x there exists a word w of length at most $n - r$ such that $|f_w^{-1}(x)| > |x|$.
- We proceed «from below» extending a singleton all the way to a maximal synchronizable set. Once such set is obtained we add another state and extend this to a union of two maximal synchronizable sets etc.

This is how we extend

Consider an automaton with 9 states. It might have the following maximal synchronizable subsets:
$(1,1,1,0,0,0,0,0,0)$, $(0,0,0,1,1,1,0,0,0)$, $(0,0,0,0,0,0,1,1,1)$.
Step 1. $(1,0,0,0,0,0,0,0,0) \xrightarrow{w_1} (1,0,1,0,0,0,0,0,0) \xrightarrow{w_2}$
$(1,1,1,0,0,0,0,0,0)$
Step 2. $(1,1,1,0,1,0,0,0,0) \xrightarrow{w_3} (0,0,0,1,0,1,1,1,1) \xrightarrow{w_4}$
$(0,0,0,1,1,1,1,1,1)$
Step 3. $(1,0,0,1,1,1,1,1,1) \xrightarrow{w_5} (1,0,1,1,1,1,1,1,1) \xrightarrow{w_6}$
$(1,1,1,1,1,1,1,1,1)$
Hence, $(1,1,1,1,1,1,1,1,1) \xrightarrow{w_6 w_5 w_4 w_3 w_2 w_1} (1,0,0,0,1,0,1,0,0)$

Lower bounds for Eulerian automata

Proposition 5

For every n and every $r < n$ such that $n = (4p+1)r$, there exists an n-state Eulerian automaton $\mathcal{A}$ of rank r with $\mathrm{mrt}(\mathcal{A}) = \frac{n^2 - r^2}{2r} - 1$.

Proposition 6

For every n and every r such that r divides n and $n/r \geq 3$ is odd, there exists an n-state binary Eulerian automaton $\mathcal{A}$ of rank r having $\mathrm{mrt}(\mathcal{A}) \geq \frac{(n-2r)^2 + nr}{2r}$.

Circular automata

- an automaton is called *circular* if there is a letter which acts on its set of states as a cyclic permutation
- in particular, the automata of Černý series $\mathcal{C}_n$ are circular
- Dubuc (in 1998) proved the Černý Conjecture for this class
- we apply results about Eulerian automata to give another proof of a quadratic upper bound on minimum rank threshold

Theorem 3

Every n-state circular automaton of rank $r < n$ has a minimum rank word of length at most $(2n - r - 1)(n - r - 1) + 1$.

Circular automata

- an automaton is called *circular* if there is a letter which acts on its set of states as a cyclic permutation
- in particular, the automata of Černý series $\mathcal{C}_n$ are circular
- Dubuc (in 1998) proved the Černý Conjecture for this class
- we apply results about Eulerian automata to give another proof of a quadratic upper bound on minimum rank threshold

Theorem 3

Every n-state circular automaton of rank $r < n$ has a minimum rank word of length at most $(2n - r - 1)(n - r - 1) + 1$.

Circular automata

- an automaton is called *circular* if there is a letter which acts on its set of states as a cyclic permutation
- in particular, the automata of Černý series $\mathcal{C}_n$ are circular
- Dubuc (in 1998) proved the Černý Conjecture for this class
- we apply results about Eulerian automata to give another proof of a quadratic upper bound on minimum rank threshold

Theorem 3

Every n-state circular automaton of rank $r < n$ has a minimum rank word of length at most $(2n - r - 1)(n - r - 1) + 1$.

Future directions

- exploring the horizons of an «averaging method»: reducing to Eulerian case, (Perron-Frobenius) weights etc.
- gaps between upper and lower bounds on mrt(·): which way should we go?

Future directions

- exploring the horizons of an «averaging method»: reducing to Eulerian case, (Perron-Frobenius) weights etc.
- gaps between upper and lower bounds on mrt(·): which way should we go?

Key references

1. Volkov M.: Synchronizing Automata and the Černý Conjecture. LATA 2008. LNCS, vol. 5196. pp. 11—27, 2008.
2. Černý, J.: Poznámka k homogénnym experimentom s konečnými automatmi. Matematicko-fyzikalny Casopis Slovensk. Akad. Vied 14(3), pp. 208–216, 1964.
3. Pin J.: Le problème de la synchronisation et la conjecture de Černý. Noncommutative structures in algebra and geometric combinatorics, Volume 109, pp. 37–48, 1981.
4. Kari J.: Synchronizing finite automata on Eulerian digraphs. Theoretical Computer Science, Volume 295 (1–3), pp. 223–232, 2003.
5. Gusev, V.V.: Lower bounds for the length of reset words in Eulerian automata. International Journal of Foundations of Computer Science 24(2), 251–262 (2013)