Journal of Automata, Languages and Combinatorics **24** (2019) 2–4, 367 – 373

AN IMPROVEMENT TO A RECENT UPPER BOUND FOR SYNCHRONIZING WORDS OF FINITE AUTOMATA

Yaroslav Shitov

Kvartira 4, Izumrudnaya ulitsa 65
129346 Moscow, Russia
yaroslav-shitov@yandex.ru

ABSTRACT

It has been known since the 60's that every synchronizing complete discrete n-state automaton admits a reset word of length at most $\alpha n^3 + o(n^3)$ for some absolute constant α. J.-E. Pin and P. Frankl proved this statement with $\alpha = 1/6 = 0.1666\ldots$ in 1982, and this bound remained best known until 2017, when M. Szykuła decreased its value to $\alpha \approx 0.1664$. In this note, we present a modification to the latest approach which leads to a more substantial improvement of $\alpha \leq 0.1654$.

Keywords: automata theory, Černý conjecture

1. Introduction

Let $\mathcal{A} = (Q, \Sigma, \delta)$ be a *deterministic finite automaton*, where Q is a finite set of *states*, Σ is a finite *alphabet*, and $\delta : Q \times \Sigma \to Q$ is a *transition function*, which assigns a mapping $Q \to Q$ to every letter of Σ. This function naturally extends to an action $Q \times \Sigma^* \to Q$ of the free monoid Σ^* on Q, and this action is still denoted by δ. For a subset $S \subseteq Q$ and a word $w \in \Sigma^*$, we define $S \cdot w$ as the set of all images $s \cdot w$ of elements $s \in S$ under the action of w. The cardinality of $Q \cdot w$ is called the *rank* of a word w; this quantity is denoted by $\mathrm{rk}_{\mathcal{A}}\, w$ or simply $\mathrm{rk}\, w$ if the choice of $\mathcal{A}$ is clear from the context. The *rank of an automaton* is defined as the smallest possible rank of a word. An automaton $\mathcal{A}$ of rank one is called *synchronizing*, and the length of the shortest rank-one words is called the *reset threshold* of $\mathcal{A}$ and denoted by $\mathrm{rt}(\mathcal{A})$.

Upper bounds on reset thresholds of synchronizing automata were a topic of extensive research in the last 50 years, and one of the main goals of this study is a famous conjecture stating that $\mathrm{rt}(\mathcal{A}) \leq (n-1)^2$ for any synchronizing n-state automaton $\mathcal{A}$; this statement was considered many years ago by different authors and became known as the *Černý conjecture* (see a historical survey in [11]). There is a lot of progress on this question for different special classes of automata [5, 7, 9], but the general version of the Černý conjecture remains wide open. The cubic upper bounds on the reset threshold, that is, inequalities of the form $\mathrm{rt}(\mathcal{A}) \leq \alpha n^3 + o(n^3)$ for some fixed α, have been known since 1966, see [6]. After a series of improvements [1, 2, 4, 5], the

progress stuck for 35 years on the celebrated $\alpha = 1/6 = 0.1666\ldots$ bound of J.-E. Pin and P. Frankl [3, 5]. In 2011, A. Trahtman [10] discovered an idea of how to find a relatively short word of rank at most $n/2$, and M. Szykuła [8] combined it with a neat linear algebraic argument and finally improved the upper bound to $\alpha \approx 0.1664$ in 2017. The purpose of this note is to modify the latter approach and get a more substantial improvement of $\alpha \leq 0.1654$.

2. Modifying the Method

From now on, we denote by $\mathcal{A} = (Q, \Sigma, \delta)$ a synchronizing automaton with n states, and we define the *corank*, or *deficiency*, of a word $w \in \Sigma^*$ as $n - \operatorname{rk} w$. Our aim is to give a relevant modification of the following theorem, which plays a crucial role in [8].

Theorem 1 [8, Lemma 2]. *Let A and S be subsets of Q satisfying $\emptyset \subsetneq A \subsetneq S$. Suppose that there is a word $w \in \Sigma^*$ such that $A \not\subseteq S \cdot w$. Then there exists a word w of length at most $n - |A|$ satisfying either*

(I) *$A \not\subseteq S \cdot w$ or*

(II) *$|S \cdot w| < |S|$.*

In [8], a successive application of Theorem 1 was used to construct a word ω that satisfies $A \not\subseteq S \cdot \omega$. In order to develop a more accurate counting approach, we need a more explicit formulation of this method.

Theorem 2. *Let $u \in \Sigma^*$ be a word of length l and corank $r \in [1, n/2 - 1]$. Assume that, for an integer λ, there exists a word v of length λ such that $\operatorname{rk} v \leq \operatorname{rk} v'$ for any word v' of length at most $\lambda + 2r$. Then there is a word of length at most $l + \lambda + 2r$ and corank at least $r + 1$.*

Proof. For a state σ in $Q \cdot u$, we denote by $\sigma \cdot u^{-1}$ the preimage of σ under the mapping $q \to q \cdot u$. Let A be the union of all those preimages $\sigma \cdot u^{-1}$ which are singleton sets; according to Lemma 7 in [8], one has $|A| \geq n - 2r$.

Now we want to find a word w of length at most $2r$ such that $A \not\subseteq Q \cdot v \cdot w$, which would allow us to find an element $a \in A$ satisfying $a \notin Q \cdot v \cdot w$, which would imply $a \cdot u \notin Q \cdot v \cdot w \cdot u$ and thus lead to a desired conclusion $Q \cdot u \supsetneq Q \cdot v \cdot w \cdot u$. Such a word w is found immediately if A and $S := Q \cdot v$ satisfy the assumptions of Theorem 1, because the second possibility of its conclusion means that $|Q \cdot v \cdot w| < |Q \cdot v|$ and contradicts the assumption of the current theorem.

As to the assumptions of Theorem 1, the one in the second sentence holds because our automaton is synchronizing. In particular, there should be a letter $b \in \Sigma$ such that $A \cdot b \neq A$. So if A was equal to $Q \cdot v$, then we could have taken $w = b$ and proceed as in the previous paragraph, and, similarly, if A was not a subset of $Q \cdot v$, we could have taken w to be the empty word and do the same thing. □

One more theorem is needed before we can proceed to counting — we cannot improve on the Pin–Frankl bound without using the Pin–Frankl bound.

Theorem 3 [3, 5]. *Let $u \in \Sigma^*$ be a word of length l and corank $r \leq n-2$. Then there is a word of length at most $l + (r+1)(r+2)/2$ and corank at least $r+1$.*

3. Counting

As Theorem 2 suggests, we are going to study the gaps between the smallest lengths of words with ranks taking consecutive pairs of values. Formally speaking, we denote by λ_i the smallest length of a word with corank at least $i \in \{0, \ldots, n-1\}$; we obviously have $0 = \lambda_0 \leq \cdots \leq \lambda_{n-1} = \mathrm{rt}(\mathcal{A})$. We also write $\lambda_n = +\infty$ and define ρ as the smallest corank satisfying $\lambda_{\rho+1} - \lambda_\rho > n$.

Observation 4. *We have $\lambda_\rho < n^2$.*

Further, we set $\delta_j = \lambda_{j+1} - \lambda_j$ for any $j \in \{0, \ldots, \rho\}$, and, for any integer $r \leq n/2$, we define the quantity s_r as the number of those $j \in \{0, \ldots, \rho\}$ which satisfy $\delta_j \in \{2r-1, 2r\}$. Let us translate Theorems 2 and 3 to this language.

Theorem 5. *Let $u \in \Sigma^*$ be a word of length l and corank $r \in [1, n/2-1]$. Then there is a word of corank at least $r+1$ and length not exceeding*

$$l + \min\left\{\frac{(r+1)(r+2)}{2},\ 2(s_1 + 2s_2 + 3s_3 + \cdots + rs_r) + 2r\right\}.$$

Proof. We are allowed to put the first argument of min by Theorem 3 immediately. Further, let us pick a word v of corank τ and length λ_τ, where τ is the minimal index for which δ_τ exceeds $2r$. The length of v does not exceed the sum of all the δ_j's not exceeding $2r$, which is at most $2(s_1+2s_2+3s_3+\cdots+rs_r)$. Also, we cannot get a word of rank less than $\mathrm{rk}\, v$ unless we take $\delta_\tau > 2r$ more letters than v has — therefore, we can apply Theorem 2 and justify the second argument of min. □

Corollary 6. *The reset threshold of $\mathcal{A}$ does not exceed*

$$\frac{7}{48}n^3 + 2\sum_{r=\rho}^{\lfloor n/2 \rfloor} \min\left\{\frac{r^2}{4}, 1s_1 + \cdots + rs_r\right\} + 3n^2.$$

Proof. We use Observation 4 to get a word of corank at least ρ and length at most n^2, then we upgrade it to a word of rank at most $\lceil n/2 \rceil$ by a successive application of Theorem 5, and then we construct a synchronizing word again by an iterative application of Theorem 3 (which takes a total of at most $7n^3/48$ additional letters). Also, the expressions under the minimum were simplified by isolating the $O(n^2)$ terms in the last summand. □

The following statement is going to complete the proof of our main result.

Proposition 7. *Let n and $\rho < n/2$ be positive integers, and let $k = \lfloor n/2 \rfloor$. Let $s_1, \dots, s_k$ be nonnegative real numbers satisfying $s_1 + \dots + s_k \le \rho$. Then*

$$\varphi(s_1, \dots, s_k) := \sum_{r=\rho}^{k} \min\left\{\frac{r^2}{4}, 1s_1 + \dots + rs_r\right\} \le \frac{15\,625\,n^3}{1\,597\,536} + o(n^3). \tag{1}$$

The numbers $s_1, \dots, s_k$ appearing in Corollary 6 are clearly nonnegative and have the sum not exceeding ρ, so we can apply Proposition 7 and get

$$\left(\frac{7}{48} + \frac{2 \cdot 15\,625}{1\,597\,536}\right) n^3 + o(n^3) \tag{B}$$

or $0.1654n^3 + O(1)$ as an upper bound for the reset threshold of $\mathcal{A}$.

4. Proving Proposition 7

The last section is devoted to a solution of the optimization problem appearing in Proposition 7. First, we restrict our attention to a certain special case.

Claim 8. *It is sufficient to prove Proposition 7 under the additional assumptions of $s_1 = \dots = s_{\rho-1} = 0$ and*

$$1s_1 + \dots + \tau s_\tau \le \tau^2/4 \quad \textit{for all } \tau \in \{\rho, \dots, k\} \tag{2_τ}$$

(where the latter says that the minimum is always attained at the second argument).

Proof. Let us define

$$s'_r = \begin{cases} 0, & \text{if } r < \rho, \\ s_r, & \text{if } r > \rho, \text{ and} \\ (1s_1 + \dots + \rho s_\rho)/\rho, & \text{if } r = \rho. \end{cases}$$

The new values are nonnegative and sum to at most $s_1 + \dots + s_k \le \rho$, so they satisfy the assumptions of Proposition 7. Also, we have $1s_1 + \dots + rs_r = 1s'_1 + \dots + rs'_r$ for all $r \ge \rho$, so the arguments of the minima do not change, and we can pass to $(s'_1, \dots, s'_k)$ without loss of generality.

Now, for a tuple $s = (s_1, \dots, s_k)$ not satisfying one of the conditions (2_τ), we define $t(s)$ as the smallest τ for which it fails and set $t := t(s)$. In other words, we have

$$\alpha := 1s_1 + \dots + ts_t > t^2/4 \quad \text{and} \quad 1s_1 + \dots + rs_r \le r^2/4 \quad \text{for all } r < t.$$

Then we set $s'_t = s_t + t/4 - \alpha/t$ and $s'_r = s_r$ for $r \notin \{t, t+1\}$, and also, if $t \ne k$, we define $s'_{t+1} = s_{t+1} - t/4 + \alpha/t$. The values s'_r are again nonnegative and sum to at most $s_1 + \dots + s_k \le \rho$, so they satisfy the assumptions of Proposition 7. Also, we have $\varphi(s') \ge \varphi(s)$ because the summands with $r \in [\rho, t]$ do not change, and the $(t+1)$-st and later summands could not have decreased. Finally, we note that, even if s' still does not satisfy some of the conditions (2_τ), we still can prove the second statement of the current theorem by induction because $t(s') > t(s)$. □

From now on, we assume that the conditions $s_1 = \cdots = s_{\rho-1} = 0$ and (2_τ) hold; we also recall that $s_r \geq 0$ and $s_1 + \cdots + s_k \leq \rho$. We call the set of all tuples $(s_1, \ldots, s_k)$ satisfying these conditions a *feasible set*; Claim 8 allows us to restrict Proposition 7 to it. The feasible set is compact (for any fixed ρ), so the function φ should have a maximum point $\sigma = (0, \ldots, 0, \sigma_\rho, \ldots, \sigma_k)$. Let β and γ be, respectively, the minimal and maximal indices i satisfying $\sigma_i \neq 0$.

Claim 9. *We have* $1\sigma_1 + \cdots + r\sigma_r = r^2/4$ *for all* $r \in \{\beta + 1, \ldots, \gamma - 1\}$.

Proof. Assume the converse and find the maximal $\nu \in \{\beta + 1, \ldots, \gamma - 1\}$ for which $1\sigma_1 + \cdots + \nu\sigma_\nu < \nu^2/4$. Then we pick a sufficiently small $\varepsilon > 0$ and define

$$\sigma'_\beta = \sigma_\beta - \varepsilon, \quad \sigma'_\nu = \sigma_\nu + \varepsilon(\nu - \beta + 1), \quad \sigma'_{\nu+1} = \sigma_{\nu+1} - \varepsilon(\nu - \beta),$$

and $\sigma'_r = \sigma_r$ for $r \notin \{\beta, \nu, \nu+1\}$. The tuple $\sigma' = (\sigma'_1, \ldots, \sigma'_k)$ sums to $\sigma_1 + \cdots + \sigma_k \leq \rho$, and its coordinates are nonnegative for a sufficiently small ε. Further, the sum as in (2_τ) could not have increased for $\tau < \nu$; the same sum but with $\tau = \nu$ has changed by something proportional to ε and so it could not have overcome $\nu^2/4$. Finally, such a sum with $\tau > \nu$ did not change as we can check directly, so the tuple σ' belongs to the feasible set. Finally, we have

$$\varphi(\sigma) = \sum_{r=\rho}^{k} (k - r + 1)\, r\, \sigma_r$$

if σ is in the feasible set, and one can get the inequality $\varphi(\sigma') > \varphi(\sigma)$, which contradicts the maximality of σ. (Such an inequality can be deduced by a straightforward computation, but let us point it out that it follows from the strict concavity of the sequence of the coefficients of $\sigma_\rho, \ldots, \sigma_k$ in the above expression for φ.) □

Now we are going to employ Claim 9 to complete the proof of Proposition 7. First, we have $r\sigma_r = r^2/4 - (r-1)^2/4$ or

$$\sigma_r = 0.5 - 0.25/r \quad \text{for all } r \in \{\beta + 2, \ldots, \gamma - 1\}. \tag{3}$$

Secondly, we have $\beta\sigma_\beta + (\beta + 1)\sigma_{\beta+1} = (\beta + 1)^2/4$ or

$$\sigma_\beta + \sigma_{\beta+1} \geq 0.25(\beta + 1). \tag{4}$$

Summing the inequalities (3) over all $r \in \{\beta + 2, \ldots, \gamma - 1\}$ and adding the inequality (4), we get at most $\sigma_1 + \cdots + \sigma_k \leq \rho \leq \beta$ on the left-hand side, or

$$\beta \geq 0.25(\beta + 1) + 0.5(\gamma - \beta - 2) - 0.25 \ln n, \tag{5}$$

where the logarithm is used to bound the harmonic numbers arising in the summation. Now we estimate $\varphi(\sigma)$ using Claim 9 again. Clearly, the summands corresponding to $r < \beta$ are zero in the definition (1) of φ, the summands with $r \in \{\beta, \ldots, \gamma\}$ cannot exceed $0.25r^2$, and the summands with $r > \gamma$ are at most $0.25\gamma^2$. We get

$$\varphi(\sigma) \leq \sum_{r=\beta}^{\gamma} 0.25r^2 + 0.25(0.5n - \gamma)\gamma^2,$$

or

$$\varphi(\sigma) \leq \psi(\beta, \gamma) := (-2\beta^3 - 4\gamma^3 + 3\gamma^2 n + 6\gamma^2 + 6\gamma + 2)/24.$$

It remains to maximize $\psi(\beta, \gamma)$ subject to $0 \leq \beta \leq \gamma \leq 0.5n$ and (5); these inequalities define a quadrilateral Δ with vertices

$$(0, 0), \quad (0, 0.5 \ln n + 1.5), \quad (0.5n, 0.5n), \quad (0.2n - 0.2 \ln n - 0.6, 0.5n).$$

Being strongly monotone in β, the function ψ cannot have a maximum inside Δ, so it remains to perform a basic calculus task and maximize ψ on the edges. The computation shows that the maximum of ψ on Δ is attained at the point $(25n/129, 125n/258) + o(n)$ and confirms that it is equal to the right-hand side of (1). Therefore, the proof of our main result is complete.

As a final remark, let us note that we were not interested to optimize the $o(n^3)$ term of our bound (B), but a direct computation of the previous paragraph gives an $O(n^2 \log n)$ estimate of this term. A more careful application of Claim 9 would make it $O(n^2)$, with explicit and reasonably small coefficients of the powers of n.

Acknowledgements

I would like to thank M. Szykuła for interesting comments and feedback on the first draft. The paper spent some time under review before getting rejected from *Journal of Combinatorial Theory, Series A*, and I am grateful to the anonymous referee of that journal for careful reading of the paper and several suggestions on its presentation. I am grateful to M. Volkov, the Guest Editor of this special issue of the *Journal of Automata, Languages and Combinatorics*, for fast processing the paper.

References

[1] J. ČERNÝ, Poznámka k homogénnym eksperimentom s konečnými automatmi. *Matematicko-fyzikalny Časopis Slovenskej Akadémie Vied* **14** (1964) 3, 208–216. (Translation: A note on homogeneous experiments with finite automata. *Journal of Automata, Languages and Combinatorics* **24** (2019) 2–4, 123–132).

[2] J. ČERNÝ, A. PIRICKÁ, B. ROSENAUEROVÁ, On directable automata. *Kybernetica* **7** (1971) 4, 289–298.

[3] P. FRANKL, An extremal problem for two families of sets. *European Journal of Combinatorics* **3** (1982), 125–127.

[4] J.-E. PIN, Sur les mots synchronisants dans un automate fini. *Elektronische Informationsverarbeitung und Kybernetik* **14** (1978) 6, 297–303.

[5] J.-E. PIN, On two combinatorial problems arising from automata theory. *Annals of Discrete Mathematics* **17** (1983), 535–548.

[6] P. H. STARKE, Eine Bemerkung über homogene Experimente. *Elektronische Informationsverarbeitung und Kybernetik (later Journal of Information Processing and Cybernetics)* **2** (1966) 4, 257–259. (Translation: A remark about homogeneous experiments. *Journal of Automata, Languages and Combinatorics* **24** (2019) 2–4, 133–137).

[7] B. Steinberg, The Černý conjecture for one-cluster automata with prime length cycle. *Theoretical Computer Science* **412** (2011) 39, 5487–5491.

[8] M. Szykuła, Improving the upper bound on the length of the shortest reset word. In: R. Niedermeier, B. Vallée (eds.), *Symposium on Theoretical Aspects of Computer Science 2018, 35th Symposium on Theoretical Aspects of Computer Science, STACS 2018, Caen, France, February 28 – March 3, 2018, Proceedings*. LIPIcs 96, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018, 56:1–56:13.

[9] A. N. Trahtman, The Černý conjecture for aperiodic automata. *Discrete Mathematics and Theoretical Computer Science* **9** (2007) 2, 3–10.

[10] A. N. Trahtman, Modifying the upper bound on the length of minimal synchronizing word. In: O. Owe, M. Steffen, J. A. Telle (eds.), *Fundamentals of Computation Theory, 18th International Symposium, FCT 2011, Oslo, Norway, August 22 – 25, 2011, Proceedings*. LNCS 6914, Springer, 2011, 173–180.

[11] M. V. Volkov, Synchronizing automata and the Černý conjecture. In: C. Martín-Vide, F. Otto, H. Fernau (eds.), *Language and Automata Theory and Applications, Second International Conference, LATA 2008, Tarragona, Spain, March 13–19, 2008, Revised Papers*. LNCS 5196, Springer, 2008, 11–27.

(Received: June 17, 2019; revised: July 25, 2019)