Exponent Sequences of Labeled Digraphs vs Reset Thresholds of Synchronizing Automata

Mikhail Volkov*

Ural Federal University, Lenina st. 51, 620083, Ekaterinburg, Russia
Mikhail.Volkov@usu.ru

A DFA $\mathscr{A} = \langle Q, \Sigma \rangle$ is called *synchronizing* if the action of some word $w \in \Sigma^*$ resets $\mathscr{A}$, that is, leaves the automaton in one particular state no matter at which state in Q it is applied: $q \cdot w = q' \cdot w$ for all $q, q' \in Q$. Any such word w is said to be a *reset word* for the DFA. The minimum length of reset words for $\mathscr{A}$ is called the *reset threshold* of $\mathscr{A}$.

In 1964 Černý [2] constructed for each $n > 1$ a synchronizing automaton $\mathscr{C}_n$ with n states whose reset length is $(n-1)^2$. Soon after that he conjectured that these automata represent the worst possible case, that is, every synchronizing automaton with n states can be reset by a word of length $(n-1)^2$. This simply looking conjecture resists researchers' efforts for more than 40 years. Even though the conjecture has been confirmed for various restricted classes of synchronizing automata, no upper bound of magnitude $O(n^2)$ for the reset threshold of n-state synchronizing automata is known in general. The best upper bound achieved so far is $\frac{n^3-n}{6}$, see [6].

One of the difficulties that one encounters when approaching the Černý conjecture is that there are only very few *extreme* automata, that is, n-state synchronizing automata with reset threshold $(n-1)^2$. In fact, the Černý series $\mathscr{C}_n$ is the only known infinite series of extreme automata. Besides that, only a few isolated examples of such automata have been found. Moreover, even *slowly* synchronizing automata, that is, automata with reset length close to the Černý bound are very rare. This empirical observation is supported also by probabilistic arguments. For instance, the probability that a composition of $2n$ random self-maps of a set of size n is a constant map tends to 1 as n goes to infinity [5]. In terms of automata, this result means that the reset threshold of a random automaton with n states and at least $2n$ input letters does not exceed $2n$. For further results of the same flavor see [7, 8]. Thus, there is no hope to

* This talk is based on a joint work of Dmitry Ananichev, Vladimir Gusev and the speaker supported by the Russian Foundation for Basic Research, grant 10-01-00524, and by the Federal Education Agency of Russia, grant 2.1.1/13995.

find new examples of slowly synchronizing automata by a lucky chance or via a random sampling experiment.

We therefore have designed and performed a set of exhaustive search experiments. A brief description of our experiments and some theoretical analysis of their outcome are presented in [1]. One of the main observations reported in [1] was a remarkable similarity between the distribution of reset thresholds of synchronizing automata and the distribution of exponents of primitive digraphs. In particular, we were able to deduce in a uniform way several series of slowly synchronizing automata, both new and already known ones, from some classical series of primitive digraphs with large exponents from [3, 9].

Despite this initial success, it turns out that the notion of exponent is too weak to be useful for isolating synchronizing automata with maximal reset threshold in some important classes, e.g. in the class of Eulerian automata. The reason for this is that we discard too much information when passing from synchronizability to primitivity—we forget anything but length about paths labeled by reset words. Thus, we have tried another approach in which more information is preserved, namely, the Parikh vectors of the paths are taken into account. Let $\mathscr{A} = \langle Q, \Sigma \rangle$ be a DFA with $|\Sigma| = k$ and fix some ordering of the letters in Σ. We define a subset $\mathrm{E}_1(\mathscr{A})$ of $\mathbb{N}_0^k$ as follows: a vector $\mathsf{v} \in \mathbb{N}_0^k$ belongs to $\mathrm{E}_1(\mathscr{A})$ if and only if there is state $r \in Q$ such that for every $p \in Q$, there exists a path from p to r such that v is the Parikh vector of the path's label. If the set $\mathrm{E}_1(\mathscr{A})$ is non-empty, then the automaton $\mathscr{A}$ is called 1-*primitive*. The minimum value of the sum $i_1 + i_2 + \cdots + i_k$ over all k-tuples $(i_1, i_2, \ldots, i_k)$ from $\mathrm{E}_1(\mathscr{A})$ is called the 1-*exponent* of $\mathscr{A}$ and denoted by $\exp_1(\mathscr{A})$. Clearly, every synchronizing automaton $\mathscr{A}$ is 1-primitive and $\exp_1(\mathscr{A})$ serves as a lower bound for the reset threshold of $\mathscr{A}$. One can find some applications of this lower bound in [4].

Here we suggest a further generalization. Let $\mathscr{A} = \langle Q, \Sigma \rangle$ be a DFA with $Q = \{1, 2, \ldots, n\}$ and let k be a non-negative integer. We say that the automaton $\mathscr{A}$ is k-*primitive* if there exist words $u_1, u_2, \ldots, u_n$ such that $1 \cdot u_1 = 2 \cdot u_2 = \cdots = n \cdot u_n$ and every word of length at most k occurs as a factor in each of $u_1, u_2, \ldots, u_n$ the same number of times. Note that the last condition implies that the words $u_1, u_2, \ldots, u_n$ have the same length. The minimal length of words that witness k-primitivity of $\mathscr{A}$ is called the k-*exponent* of $\mathscr{A}$ and is denoted by $\exp_k(\mathscr{A})$.

Consider now an arbitrary synchronizing automaton $\mathscr{A}$. It is clear that $\mathscr{A}$ is k-primitive for every k and $\exp_k(\mathscr{A})$ serves as a lower bound for the reset threshold of $\mathscr{A}$. Thus, we have the following non-decreasing sequence:

$$\exp_1(\mathscr{A}) \leq \cdots \leq \exp_k(\mathscr{A}) \leq \exp_{k+1}(\mathscr{A}) \leq \ldots . \quad (1)$$

At every next step we require that words $u_1, u_2, \ldots, u_n$ get more similar to each other than they were in previous step. Thus, sooner or later these words "converge" to a reset word and the sequence stabilizes at the reset threshold of $\mathscr{A}$. Our hope is that studying the sequence (1) may shed new light on the Černý conjecture.

References

1. Ananichev, D.S., Gusev, V.V., Volkov, M.V.: Slowly synchronizing automata and digraphs. In: Hliněný, P., Kučera, A. (eds.), Mathematical Foundations of Computer Science, Lect. Notes Comp. Sci., vol. 6281, pp. 55–65. Springer, Heidelberg (2010)
2. Černý, J.: Poznámka k homogénnym eksperimentom s konečnými automatami. Matematicko-fyzikalny Časopis Slovensk. Akad. Vied 14(3) 208–216 (1964) (in Slovak)
3. Dulmage, A.L., Mendelsohn, N.S.: Gaps in the exponent set of primitive matrices. Ill. J. Math. 8, 642–656 (1964)
4. Gusev, V.V.: Lower bounds for the length of reset words in Eulerian automata. In: 5th Workshop on Reachability Problems (accepted)
5. Higgins, P.M.: The range order of a product of i transformations from a finite full transformation semigroup, Semigroup Forum 37, 31–36 (1988)
6. Pin, J.-E.: On two combinatorial problems arising from automata theory. Ann. Discrete Math. 17, 535–548 (1983)
7. Skvortsov, E. Tipikin, E.: Experimental study of the shortest reset word of random automata. In: Bouchou-Markhoff, B. et al. (eds.), Implementation and Application of Automata. Lect. Notes Comp. Sci., vol. 6807, pp. 290–298. Springer, Heidelberg (2011)
8. Skvortsov, E., Zaks, Yu.: Synchronizing random automata. Discr. Math. Theor. Comput. Sci. 12(4), 95–108 (2010)
9. Wielandt, H.: Unzerlegbare, nicht negative Matrizen. Math. Z. 52, 642–648 (1950) (in German)