

# Experiments with Synchronizing Automata

Andrzej Kisielewicz\*, Jakub Kowalski\*\*, and Marek Szykuła\*\*\*

Department of Mathematics and Computer Science, University of Wrocław  
andrzej.kisielewicz@math.uni.wroc.pl, jko@cs.uni.wroc.pl,  
msz@cs.uni.wroc.pl

**Abstract.** We have improved an algorithm generating synchronizing automata with a large length of the shortest reset words. This has been done by refining some known results concerning bounds on the reset length. Our improvements make possible to consider a number of conjectures and open questions concerning synchronizing automata, checking them for automata with a small number of states and discussing the results. In particular, we have verified the Černý conjecture for all binary automata with at most 12 states, and all ternary automata with at most 8 states.

## 1 Introduction

A deterministic finite automaton  $\mathcal{A}$  is  $\langle Q, \Sigma, \delta \rangle$ , where  $Q$  is the set of the states,  $\Sigma$  is the input alphabet, and  $\delta: Q \times \Sigma \rightarrow Q$  is the (complete) transition function. Throughout the paper, by  $n$  we denote the number of states  $|Q|$ . If  $|\Sigma| = k$  then  $\mathcal{A}$  is called  $k$ -ary. The transition function  $\delta$  is naturally extended to a function  $2^Q \times \Sigma^* \rightarrow 2^Q$ . The *image* of  $S \subseteq Q$  under the action of a word  $w \in \Sigma^*$  is  $Sw = \{\delta(q, w) \mid q \in S\}$ . The *rank* of a word  $w \in \Sigma^*$  is  $|Qw|$ , and the *rank* of  $\mathcal{A}$  is the minimal rank of a word over  $\mathcal{A}$ . For a non-empty subset  $\Sigma' \subseteq \Sigma$ , we may define the automaton  $\mathcal{A}' = \langle Q, \Sigma', \delta' \rangle$ , where  $\delta'$  is the natural restriction of  $\delta$  to  $\Sigma'$ . In such a case  $\mathcal{A}$  is called an *extension* of  $\mathcal{A}'$ . The automata of rank 1 are called *synchronizing*, and each word  $w$  with  $|Qw| = 1$  is called a *synchronizing* (or *reset*) word for  $\mathcal{A}$ . An automaton is *irreducibly synchronizing* if it is not an extension of a synchronizing automaton over a smaller alphabet.

We are interested in the length of a shortest reset word for  $\mathcal{A}$  (there may be more than one word of the same shortest length). We call it the *reset length* of  $\mathcal{A}$ . The famous Černý conjecture states that every synchronizing automaton  $\mathcal{A}$  with  $n$  states has a reset word of length  $\leq (n - 1)^2$  [9]. This conjecture was formulated by Černý in 1969 and is considered the longest-standing open problem in combinatorial theory of finite automata. So far, the conjecture has

---

\* Supported in part by the National Science Centre, Poland under project number 2012/07/B/ST1/03318.

\*\* Supported in part by the National Science Centre, Poland under project number 2015/17/B/ST6/01893.

\*\*\* Supported in part by the National Science Centre, Poland under project number 2013/09/N/ST6/01194.

been proved only for a few special classes of automata, and a cubic upper bound  $(n^3 - n)/6 - 1$  [19] has been established, which was not improved for over 30 years (see [15, 24] for excellent surveys). The bound  $(n - 1)^2$  is met for every  $n$  by the Černý automata [9], which is the only known infinite series of automata meeting this bound (besides that, there are 8 known particular examples with  $n \leq 6$  states [22] also meeting the bound).

There were several efforts to check computationally the conjecture for all automata with a small number of states. In particular, Ananichev, Gusev, and Volkov [3, 4] have checked all binary automata with at most  $n = 9$  states, and the checking for all automata with at most  $n = 10$  states was reported in [22]. In [16], using a dedicated algorithm, we have verified the conjecture for all binary automata with  $n \leq 11$  states.

In this paper, first we describe improvements to our algorithm from [16], which are aimed at making possible verifying the conjectures for larger automata. While these are results of a rather technical nature, and may be not very interesting from theoretical point of view, they make possible to restrict the computation process to much smaller class of relevant automata, and thus to consider also automata with a larger number of states.

We extend verification of the Černý conjecture up to 12 states and present an extensive experimental study on important problems and conjectures closely related to upper bounds on reset lengths. We consider known conjectures, and restate or state new ones basing on our experiments. Most of them imply an improvement for the general cubic bound, and hence are very hard but stand as possible ways to attack the main problem. All of the conjectures are experimentally confirmed for automata with a small number of states and/or letters.

## 2 Reset Lengths of Extensions

In this section we describe two theoretical results we apply in the improved algorithm. We are interested mainly in estimating the reset length of synchronizing automata that arise as extensions of non-synchronizing automata by one letter. In some cases, we are able to provide better upper bounds than the general bound  $(n^3 - n)/6 - 1$  [19].

In particular, we search for synchronizing automata with relatively large reset length. We improve the algorithm from [16] which takes a set of  $(k - 1)$ -ary automata with  $n$  states and generates all their nonisomorphic one-letter extensions. To perform an exhaustive search over the  $k$ -ary automata with  $n$  states with some property, we need to progressively run the algorithm  $k - 1$  times starting from the complete set of non-isomorphic unary automata. However, in each run, if we know that any extension of an automaton  $\mathcal{A}$  cannot have the desired property, we can safely drop  $\mathcal{A}$  from further computations. Since the number of generated automata grows rapidly, suitable knowledge saves a lot of computational time and extends the class of the automata investigated. The technical details of the algorithm and proofs can be found in [17].

A subset  $M \subseteq Q$  of the states is called *compressible*, if there is a word  $w$  such that  $|Mw| < |M|$ . Let  $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$  be a finite automaton. We say that a sequence  $(M_i, x_i, y_i)$ ,  $(1 \leq i \leq \ell)$  of  $m$ -subsets (subsets of size  $m$ )  $M_i$  of  $Q$  and pairs of states  $x_i, y_i \in Q$  is an  *$m$ -subset Frankl-Pin sequence* if the following conditions are satisfied

1.  $x_i, y_i \in M_i$  for  $1 \leq i \leq \ell$ ;
2. either  $x_i$  or  $y_i$  is not in  $M_j$  for all  $1 \leq j < i \leq \ell$ .

If all the pairs  $\{x_i, y_i\}$  belong to a set  $P$  of pairs, we will say that this sequence is *over*  $P$ . Given a set  $P$  of *compressible* pairs, by the *synchronizing height*  $h(P)$  of  $P$  we mean the minimal  $h$  such that for each pair  $\{x, y\} \in P$  there exists a word  $w$  of length  $h$  such that  $xw = yw$ .

It is known that a shortest word compressing  $M$  cannot be longer than the length of the Frankl-Pin sequence starting from  $M$  [12] (this, in fact, is used to obtain the bound  $(n^3 - n)/6$  mentioned above). Our first technical improvement is that if the synchronizing height is smaller than the maximal length of a Frankl-Pin sequence over  $P$ , then we have

**Theorem 1.** *Let  $P$  be a set of compressible pairs in  $\mathcal{A}$ ,  $h(P)$  the synchronizing height of  $P$ , and  $p(P)$  the maximal length of a Frankl-Pin sequence over  $P$ . Then, for every compressible  $m$ -subset  $M$  of  $Q$  ( $2 \leq m \leq n$ ), there is a word compressing  $M$  whose length does not exceed*

$$\binom{n - m + 2}{2} - p(P) + h(P).$$

This result improves the estimation in [12] by the negative summand  $(p(P) - h(P))$ . It is to be combined and compared with the result by J.-E. Pin [18] saying that if  $w$  is a word of rank  $r$  and there exists a word of rank  $\leq r - 1$ , then there is such a word of length  $\leq 2|w| + n - r + 1$ . There are other results of this kind that can be used for providing bounds for extensions, as that in [6]. Unfortunately, for small values of  $n$  that are within our considerations, this does not overcome the bound from Theorem 1.

Recall that an automaton  $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$  is *one-cluster*, if it has a letter  $a \in \Sigma$  such that for every pair  $q, s \in Q$  there are  $i, j \geq 1$  such that  $qa^i = sa^j$ . This means that the graph of the transformation induced by  $a$  is connected. In particular, it has a unique cycle  $C \subseteq Q$  with the property  $Ca^i = C$  for every  $i \geq 0$ , and there is  $\ell \geq 0$  such that  $Qa^\ell = C$ . The least such  $\ell$  is called the *level* of  $\mathcal{A}$ . Steinberg [21] proved that if the length  $m$  of the cycle is prime, then the one-cluster automaton  $\mathcal{A}$  has a reset word of length at most

$$n - m + 1 + 2\ell + (m - 2)(n + \ell). \quad (1)$$

We generalize this result to arbitrary lengths and get an additional negative summand. We refine the proof of Steinberg [21] and the summand is expressed in algebraic terms of the proof. Therefore, to present the result we have to recall basic notations from [21].

Given a one cluster automaton with the notation as above, we consider the matrix representation  $\pi: \Sigma^* \rightarrow M_n(\mathbb{Q})$  defined by  $\pi(w)_{q,r} = 1$  if  $qw = r$ , and 0, otherwise. Given  $S \subseteq Q$  we define  $[S]$  to be the characteristic row vector of  $S$  in  $\mathbb{Q}^n$ ,  $[S]^T$  its transpose, and  $\gamma_S = [S]^T - (|S|/|C|)[Q]^T$ . By  $w\gamma_S$  we denote the product of corresponding matrices; in particular, the word  $w$  represents the matrix  $\pi(w)$ , and the product is a vector in the space  $\mathbb{Q}^n$ . We consider the subspace  $W_S = \text{Span}\{a^{\ell+j}\gamma_S \in \mathbb{Q}^n \mid 0 \leq j \leq m-1\}$  (cf. [21]), and the *cyclic period*  $q_S$  of  $S$ , understood as the least number  $q$  such that  $Sa^q = S$ . Now, we define  $D^*(m, k)$  to be the minimal value of  $m - q_S + \dim W_S$  taken over all vectors  $S$  with  $|S| = k$ . Then we prove the following:

**Theorem 2.** *Let  $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$  be a synchronizing automaton with  $n$  states, such that there exists a word  $w$  of length  $s$  inducing a one-cluster transformation with level  $\ell$  and cycle  $C$  of length  $m > 1$ . Then  $\mathcal{A}$  has a reset word of length at most*

$$s(\ell + m - 2)(m - 1) + (n + 1)(m - 1) + s\ell - \sum_{k=1}^{m-1} D^*(m, k).$$

One can demonstrate that this results generalizes and improves earlier bounds in [8, 21], and a careful estimation of the summand  $D^*(m, k)$  yields the currently best general bound for reset lengths of one-cluster automata:

**Corollary 3.** *A synchronizing one-cluster automaton  $\mathcal{A}$  with  $n$  states and the cycle of length  $m$  has a reset word of length at most*

$$2nm - 4m \ln \frac{m+3}{2} + 2m - n + 1 \quad (2)$$

Nevertheless, for small values of  $m$  we can compute the exact values of  $D^*(m, k)$ , and this yields considerably better bounds than the general estimation above.

### 3 Experiments and Conjectures

In this section we discuss the results of our experiments with the improved algorithm concerning various conjectures and open problems in the area.

**The Černý conjecture.** We have verified the Černý conjecture for several cases. In particular, we confirmed it for all binary automata with  $n \leq 12$  states, and for all ternary automata with  $n \leq 8$ .

Verifying the Černý conjecture for binary automata with  $n = 12$  states was the most difficult computation that we have performed here. The total time of a single processor core spent for this computation was about 100 years. We performed this on a grid in parallel using mostly about 200 cores of Quad-Core AMD Opteron(tm) Processor 8350, 2.0 GHz. The total number of automata generated by our algorithm in this case was about  $10^{15}$ .

For ternary automata with  $n = 8$  states the computation took 1.25 years of a single processor core, and we had to generate and check about  $2.1 \times 10^{10}$  automata. One may compare these numbers with the numbers of non-isomorphic initially connected automata that one would need to generate applying the technique described in [3]. The corresponding numbers are: about  $2.2 \times 10^{17}$  for binary automata with  $n = 12$  states, and  $5.7 \times 10^{17}$  for ternary automata with  $n = 8$  states.

Within the range we have considered, the only automata meeting the bound  $(n - 1)^2$  other than the Černý series are known examples with  $n \leq 6$  states that were presented in [22].

**Slowly synchronizing automata.** For the case of binary automata  $n = 12$  states, we have obtained also the complete list of strongly connected synchronizing automata with reset length  $\geq 94$ .

**Table 1.** The numbers of all non-isomorphic strongly connected synchronizing binary automata with 12 states with reset length  $\geq 94$ .

Reset length	94	95–98	99	100	101	102	103–109	110	111	112	113–120	121
Number of automata	3	0	3	21	9	2	0	2	1	1	0	1
Series						$\mathcal{H}_n, \mathcal{K}_n$		$\mathcal{E}_n, \mathcal{D}'_n$	$\mathcal{W}_n$	$\mathcal{D}'_n$		$\mathcal{C}_n$

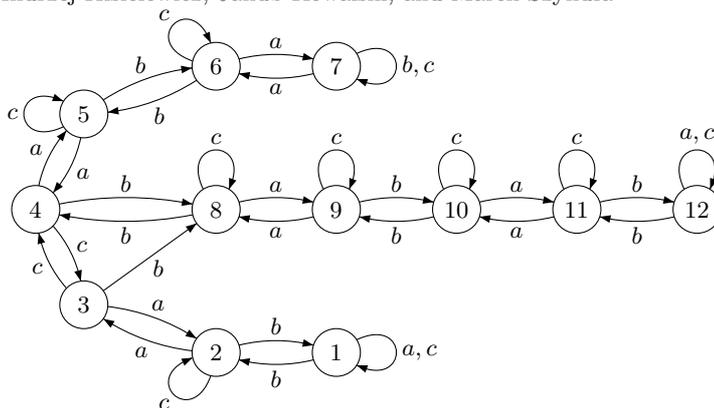
Table 1 shows the exact numbers of automata in this range, and the corresponding series according to naming from [3, 4, 16]. Here, all automata with reset length  $\geq 99$  has a similar structure of one long cycle and a small gadget (cf. [4]), and they can be generalized to series of length  $n^2 - O(n)$  as well. We confirm, for  $n \leq 12$ , [4, Conjecture 1], which is a generalization of the Černý conjecture, describing all binary synchronizing automata with reset length  $\geq n^2 - 4n + 8$  (104 for  $n = 12$ ) and stating that up to isomorphism this list is complete.

As observed in [4, 16, 22], there are gaps in the set of possible reset lengths near the Černý bound  $(n - 1)^2$ . We confirm for binary automata that for  $n = 6, 7, 8$  there is one gap, for  $n = 9, 10$  there are two gaps, and for  $n = 11, 12$  there are three gaps.

There is no binary strongly connected automaton with 12 states and reset length 95, but we have constructed such an automaton over a ternary alphabet (Fig. 1). Similarly, we know an automaton for  $n = 9$  with reset length 53 (second gap), and for  $n = 11$  with reset length 79 (third gap). This shows that the gaps, except the first one, are not necessarily preserved over larger alphabets.

**Extending words in one-cluster automata.** One-cluster automata are an important class of synchronizing automata for which a quadratic bound on reset length has been found [5, 21].

Despite several attempts [5, 8, 10, 20, 21] at improving the bounds, so far, the Černý conjecture has been proved only for one-cluster automata with a cycle of



**Fig. 1.** An irreducibly synchronizing strongly connected ternary automaton with 12 states and reset length 95.

length  $n$  (circular automata) or with a prime-length cycle. In [21] an algebraic argument making use of ascending chain of linear subspaces and averaging trick has been applied. The proof is based on the claim that any subset  $S \subset C$  on the cycle  $C$  can be extended on this cycle by a word of length at most  $\ell + n$  (we apply here the notation of Section 2). It is demonstrated that this holds in the case of prime length of  $C$ . Proving it for non-prime lengths would provide the proof of the Černý conjecture for the whole class of one-cluster automata.

We have exhaustively searched for small examples of one-cluster automata with a non-prime cycle length such that the length  $\ell + n$  of extending words is exceeded for some subset  $S$ , but found out that  $\ell + n$  is sufficient in all tested cases, instead of the value  $n + \ell + |C| - D^*(|C|, |S|)$  used to prove the bound from Theorem 2. Also, we found out that we can always use an extending word of the form  $wa^\ell$  with  $|w| \leq n$ , which is the form used in the proof for prime  $|C|$ .

**Conjecture 1.** Let  $\mathcal{A}$  be a one-cluster synchronizing automaton with a one-cluster letter  $a$  with the cycle  $C$  and level  $\ell$ . For any non-empty proper subset  $S \subset C$  there is a word  $w$  such that  $|S(wa^\ell)^{-1} \cap C| > |S|$  and  $|w| \leq n$ .

In all the cases tested, for any  $\ell$ , non-prime  $|C| < n$ , and  $|S|$  with  $1 \leq |S| < |C|$ , we found an automaton for which we needed a word  $w$  of length exactly  $n$ . So, it seems that the bound  $|w| \leq n$  is tight.

**Worst cases for the greedy compressing algorithm.** The *greedy compressing algorithm* is a well known approach for finding a reset word [11, 19, 24]. It starts from  $S = Q$ , and iteratively finds a shortest word  $w$  such that  $|Sw| < |S|$  and uses  $Sw$  for next iteration, until  $|S| = 1$ . The concatenated words  $w$  form the found reset word. The length of the resulted reset word can vary, since there is ambiguity in selection of shortest words  $w$ . By bounding the length of the found reset word we also obtain an upper bound for the reset length, and in fact, the upper bound  $(n^3 - n)/6$  for the reset length is obtained by bounding the lengths of words  $w$  for  $|S| = 2, \dots, n$  and summing these bounds [19]. It is known that

this algorithm finds a word of length  $\Omega(n^2 \log n)$  for the Černý automaton [15], but it was not clear whether it is the worst case example.

We experimentally tested the greedy algorithm for the worst cases. Here, we restricted the studied class to irreducibly synchronizing automata, as otherwise we would get a lot of trivial examples derived from automata over a smaller alphabet. By the *worst case length* we mean the maximum length of the found word by the algorithm over all selections of shortest compressing words that can be taken by the algorithm. For example, for automaton  $\mathcal{G}_1$  from Fig. 2, the worst case length is 19 and a sequence of subsets considered by the greedy algorithm in the worst case can be the following:

$$Q = \{1, 2, 3, 4, 5\} \xrightarrow{b} \{1, 2, 4, 5\} \xrightarrow{aca} \{3, 4, 5\} \xrightarrow{bcba\check{c}b} \{1, 4\} \xrightarrow{acbbcbaca} \{3\}.$$

While this requires potentially very expensive computation, the worst case length can be computed by a kind of dynamic algorithm and  $n - 1$  iterations of breadth-first search in the power automaton.

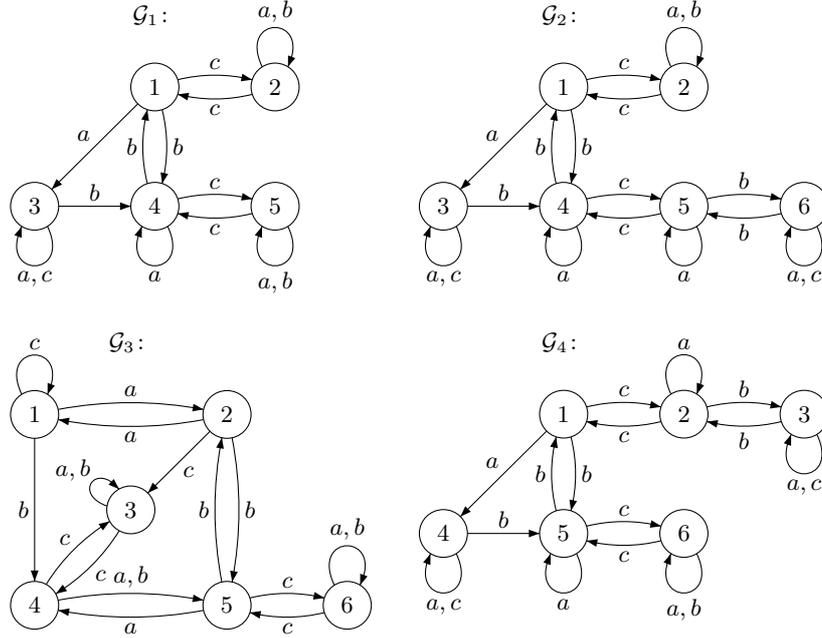
It may be surprising that the Černý automata generally do not exhibit the worst case length. We have observed that for some values  $n \geq 10$  the slowly synchronizing series  $\mathcal{W}_n$ ,  $\mathcal{D}_n''$ , and  $\mathcal{G}_n$  (see [3, 4]) exceed the worst case length of the Černý automaton with the same number of states. In addition, we have found out four particular ternary examples shown in Fig. 2 exceeding the worst case length of the Černý automaton with the same number of states, which do not seem generalizable to series. Up to isomorphism, there are no more such examples within the range we have considered (Table 2).

The results we have collected do not allow to state a reasonable conjecture. So far,  $\mathcal{W}_n$  is the best candidate for the largest worst case lengths for  $n \geq 10$ , and the Černý automata for  $n \leq 9$ , except  $\mathcal{G}_1$  and  $\mathcal{G}_2$  from Fig. 2 for  $n = 5, 6$ .

**Problem 2.** What are the largest worst case lengths of the greedy compressing algorithm of automata with  $n$  states?

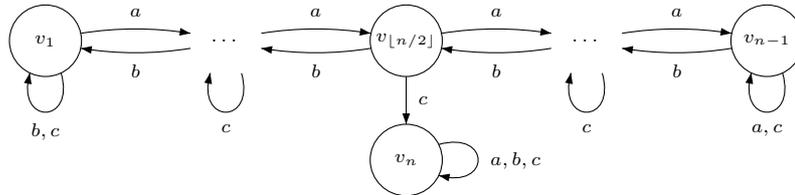
It is noticeable that the dual *greedy extending algorithm*, which starts from a singleton and uses shortest extending words rather than compressing ones, seem to have generally larger worst case lengths. For example, for the case of binary  $n = 7$  in the worst case it can find a reset word of length 48 for some strongly connected automaton, whereas the greedy compressing algorithm finds a word of length at most 43.

**Aperiodic synchronizing automata.** Recall that an automaton is *aperiodic* if there is no word inducing a transformation with a cycle of length  $\geq 2$  (the transition semigroup has only trivial subgroups). In [24] Volkov mentioned that although a quadratic upper bound for the reset length of aperiodic synchronizing automata has been proved, the largest reset length for known aperiodic automata does not exceed  $n + \lfloor n/2 \rfloor - 2$ . This length is reached by a series of binary automata constructed by Ananichev [2]. In this connection, it may be interesting to know that the same bound is also reached for every  $n > 1$  by a series of



**Fig. 2.** Automata  $\mathcal{G}_1$ ,  $\mathcal{G}_2$ ,  $\mathcal{G}_3$ , and  $\mathcal{G}_4$ , with the worst case length 19, 30, 28, and 28, and reset lengths 15, 22, 20, and 20, respectively.

irreducibly ternary aperiodic automata. It has a quite simple definition and an easy proof for the reset length (comparing with [2]). Let  $\mathcal{A}_n = \langle Q, \{a, b, c\}, \delta \rangle$ , where  $Q = \{v_1, \dots, v_n\}$ ,  $\delta(v_i, a) = v_{i+1}$  for  $1 \leq i \leq n - 2$ ,  $\delta(v_i, b) = v_{i-1}$  for  $2 \leq i \leq n - 1$ ,  $\delta(v_{\lfloor n/2 \rfloor}, c) = v_n$ , and  $\delta(v_i, x) = v_i$ , otherwise ( $x \in \Sigma$ ) (shown in Fig. 3).



**Fig. 3.** A ternary irreducibly synchronizing  $n$ -state aperiodic automaton with the reset length  $n + \lfloor n/2 \rfloor - 2$ .

Volkov<sup>1</sup> has also pointed out that  $n - 1$  may be an upper bound for the reset length in the class of *strongly connected* synchronizing aperiodic automata, but there was not enough evidence. The bound can be met trivially if the underlying digraph of the automaton is a *bidirectional path*:  $Q = 1, \dots, n$ , for every  $1 \leq i \leq$

<sup>1</sup> personal communication

$n - 1$  there are the directed edges  $(i, i + 1)$  and  $(i + 1, i)$ , and every edge that is not a loop is of that form.

Since our verifications involve a huge number of aperiodic automata, we experimentally support the following conjectures:

**Conjecture 3 (cf. [24]).** Every synchronizing aperiodic automaton with  $n > 1$  states has a reset word of length at most  $\leq n + \lceil n/2 \rceil - 2$ .

**Conjecture 4 (Volkov).** Every strongly connected synchronizing automaton has a reset word of length at most  $n - 1$ . Moreover, if this bound is met, then the underlying digraph of the automaton is a bidirectional path.

**Avoiding states.** In a recent short note [13] the authors state the following problem related to the recent unsuccessful attempt of improving the general upper bound on reset length [23]: Given a strongly connected synchronizing automaton, what is the minimal length  $\ell$  such that for any  $q \in Q$  there is a word  $w$  of length  $\leq \ell$  and such that  $q \notin Qw$ . If  $\ell \in O(n)$ , then we would obtain a better upper bound than  $(n^3 - n)/6$ .

Experimentally, we have found out what the value of  $\ell$  for a given  $n$  might be, and provided support for the following conjecture:

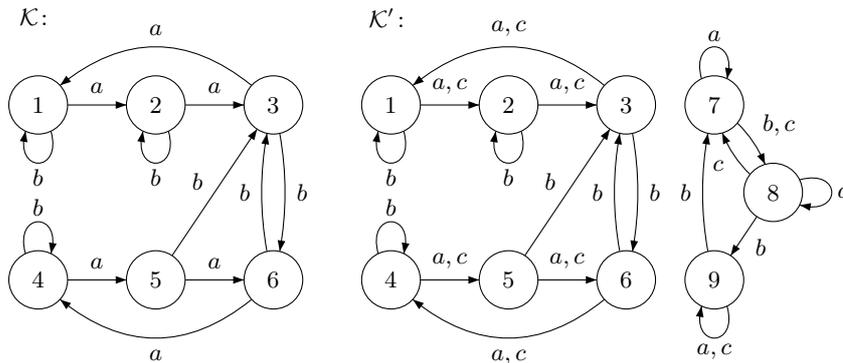
**Conjecture 5.** In a synchronizing strongly connected automaton, for any  $q \in Q$  there is a word  $w$  of length  $\leq 2n - 2$  and such that  $q \notin Qw$ . This bound is tight for  $n \geq 4$  over a ternary alphabet.

Recently, Vojtěch Vorel<sup>2</sup> discovered an infinite series of binary automata whose minimal length in question is  $2n - 4$ , which is currently the best theoretical lower bound for the problem.

**New rank conjecture.** Pin [19] proposed the following generalization of the Černý conjecture: For every  $0 < d, n$ , if there is a word of rank  $\leq n - d$ , then there is such a word of length  $\leq d^2$ . Pin proved this for  $d \leq 3$ . However, Kari [14] found a celebrated counterexample to this conjecture for  $d = 4$ , which is a binary automaton  $\mathcal{K}$  with 6 states (Fig. 4). As a consequence, a modification of this generalized conjecture was proposed restricting it to  $d$  being the rank of the considered automaton (see for example [1]). However this seems to be a quite radical restriction.

In our computations, we have found no other counterexample to Pin's conjecture except for trivial extensions and modifications. This may suggest that Kari construction works due to the number of involved states small enough, and is, in fact, an exception. By a *trivial extension* of an automaton over alphabet  $\Sigma$  we mean one obtained by adding letters to  $\Sigma$  that acts either as the identity transformation or as any letter in  $\Sigma$ . So a trivial extension has the same number of the states and the transition semigroup, and trivial extensions of the Kari automaton  $\mathcal{K}$  are counterexamples to the Pin's conjecture, for  $d = 4$ , as well. By

<sup>2</sup> personal communication, unpublished



**Fig. 4.** The Kari automaton  $\mathcal{K}$  [14], and a Kari-like automaton  $\mathcal{K}'$ .

a *disjoint union* of two automata  $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$  and  $\mathcal{A}' = \langle Q', \Sigma, \delta' \rangle$  we mean the construction where the automata have the same alphabet  $\Sigma$ , and disjoint sets of states  $Q, Q'$ , and the union is simply  $\mathcal{A} = \langle Q \cup Q', \Sigma, \delta \cup \delta' \rangle$ . If we take a disjoint union of  $\mathcal{K}$  with any *permutation automaton* (one whose letters act like permutations, or in other words, one of rank equal to its size), then again we get a counterexample to the Pin's conjecture, for  $d = 4$ . Yet, in all these automata the failure is caused by the same Kari construction on the set of the 6 states. In our experiments, we have discovered no other counterexample. This may be treated as an evidence for the conjecture we state below.

Consider the smallest class of automata containing  $\mathcal{K}$  and closed on taking trivial extension and disjoint union with permutation automata. Let us call automata in this class *Kari-like* automata (see Fig. 4). Then we have

**Conjecture 6.** For every  $d$ , if an automaton  $\mathcal{A}$  has a word of rank at most  $n - d$ , then there is such a word of length at most  $d^2$ , unless  $\mathcal{A}$  is a Kari-like automaton and  $d = 4$  (in which case there is a word of rank  $n - 4$  of length  $d^2 + 1 = 17$ ).

**Subset synchronization.** The last conjecture was posed by Ângela Cardoso:

**Conjecture 7 (Cardoso [7]).** In a synchronizing automaton, for any subset  $S$  of states there is a word  $w$  with  $|Sw| = 1$  of length at most

$$(n - 1)^2 - \left\lceil \frac{n - |S|}{|S|} \right\rceil \left( 2n - |S| \left\lceil \frac{n}{|S|} \right\rceil - 1 \right).$$

This is another generalization of the Černý conjecture, and it can be viewed as a counterpart for the rank conjecture, where we bound the length of words compressing  $Q$  to a subset of the given size, rather than a subset to a singleton.

Conjecture 7 has been proved for several special classes of automata, and the formula is tight for any subset size in the Černý series. Besides confirmation for small automata, we identified 18 particular examples of irreducibly synchronizing

automata with  $n \in \{3, 4, 5, 6\}$  states meeting the bound for some subset  $S$  that are not isomorphic to the Černý automata. Note that the conjecture is not true in general for non-synchronizing automata, as Vorel [25] has constructed a series of non-synchronizing strongly connected binary automata with subsets whose shortest synchronizing words are of exponential length.

### 3.1 Summary

Table 2 summarizes the ranges for which the discussed conjectures have been confirmed or the problems checked. The ranges vary due to different numbers of automata that have to be checked, computational complexity of verification for a single automaton, and computation time devoted for each of the problems.

**Table 2.** Experimental verification of conjectures. The numbers denote the size of the alphabet up to which the given conjecture has been checked. The symbol  $\infty$  denotes that the problem has been verified for all automata with the given number of states and any number of letters.

Problem	Number of states $n$								
	$\leq 4$	5	6	7	8	9	10	11	12
Černý conjecture and [4, Conjecture 1]	$\infty$	$\infty$	6	4	3	2	2	2	2
Conjecture 1 (one-cluster)	$\infty$	5	4	3	2	2	2		
Problem 2 (greedy algorithm)	$\infty$	6	4	3	2	2	2		
Conjecture 3 (aperiodic)	$\infty$	5	3	3	2	2	2		
Conjecture 4 (strongly connected aperiodic)	$\infty$	8	5	3	2	2	2	2	
Conjecture 5 (avoiding states)	$\infty$	8	4	3	2	2	2		
Conjecture 6 (new rank conjecture)	$\infty$	$\infty$	5	3	3	2	2		
Conjecture 7 (subset synchronization)	$\infty$	$\infty$	5	4	3	2	2		

**Acknowledgments.** We thank Mikhail Volkov for suggesting Conjecture 4, and Mikhail Berlinkov for observing that the bound for one-cluster automata can be improved for periodic subsets on the cycle, which led to an improvement of our algorithm. We thank also Vojtěch Vorel for discussing the problem of avoiding states and sharing the series. The main part of the computations was performed on a grid that belongs to Institute of Computer Science of Jagiellonian University (thanks to Adam Roman).

### References

1. J. Almeida and B. Steinberg. Matrix Mortality and the Černý-Pin Conjecture. In *DLT*, volume 5583 of *LNCS*, pp. 67–80. 2009.
2. D. S. Ananichev. The Mortality Threshold for Partially Monotonic Automata. In *DLT*, volume 3572 of *LNCS*, pp. 112–121. 2005.

3. D. S. Ananichev, V. V. Gusev, and M. V. Volkov. Slowly synchronizing automata and digraphs. In *MFCS*, volume 6281 of *LNCS*, pp. 55–65. 2010.
4. D. S. Ananichev, M. V. Volkov, and V. V. Gusev. Primitive digraphs with large exponents and slowly synchronizing automata. *Journal of Mathematical Sciences*, 192(3):263–278, 2013.
5. M.-P. Béal, M. V. Berlinkov, and D. Perrin. A quadratic upper bound on the size of a synchronizing word in one-cluster automata. *International Journal of Foundations of Computer Science*, 22(2):277–288, 2011.
6. M. Berlinkov and M. Szykuła. Algebraic Synchronization Criterion and Computing Reset Words. In *MFCS*, volume 9234 of *LNCS*, pp. 103–115. 2015.
7. Á. Cardoso. *The Černý Conjecture and Other Synchronization Problems*. PhD thesis, University of Porto, Portugal, 2014.
8. A. Carpi and F. D’Alessandro. Independent sets of words and the synchronization problem. *Advances in Applied Mathematics*, 50(3):339–355, 2013.
9. J. Černý. Poznámka k homogénnym experimentom s konečnými automatami. *Matematicko-fyzikálny Časopis Slovenskej Akadémie Vied*, 14(3):208–216, 1964.
10. L. Dubuc. Sur les automates circulaires et la conjecture de Černý. *Informatique théorique et applications*, 32:21–34, 1998.
11. D. Eppstein. Reset sequences for monotonic automata. *SIAM Journal on Computing*, 19:500–510, 1990.
12. P. Frankl. An extremal problem for two families of sets. *European Journal of Combinatorics*, 3:125–127, 1982.
13. F. Gonze, R. M. Jungers, and A. N. Trahtman. A Note on a Recent Attempt to Improve the Pin-Frankl Bound. *Discrete Mathematics and Theoretical Computer Science*, 17(1):307–308, 2015.
14. J. Kari. A counter example to a conjecture concerning synchronizing word in finite. *EATCS Bulletin*, 73:146–147, 2001.
15. J. Kari and M. V. Volkov. Černý’s conjecture and the road coloring problem. In *Handbook of Automata*. European Science Foundation, 2013.
16. A. Kisielewicz and M. Szykuła. Generating Small Automata and the Černý Conjecture. In *CIAA*, volume 7982 of *LNCS*, pp. 340–348. 2013.
17. A. Kisielewicz and M. Szykuła. Generating Synchronizing Automata with Large Reset Lengths. <http://arxiv.org/abs/1404.3311>, 2016.
18. J.-E. Pin. Utilisation de l’algèbre linéaire en théorie des automates. In *Actes du 1er Colloque AFCET-SMF de Mathématiques Appliquées II*, pp. 85–92, 1978.
19. J.-E. Pin. On two combinatorial problems arising from automata theory. In *Proceedings of the International Colloquium on Graph Theory and Combinatorics*, volume 75 of *North-Holland Mathematics Studies*, pp. 535–548, 1983.
20. B. Steinberg. The averaging trick and the Černý conjecture. *International Journal of Foundations of Computer Science*, 22(7):1697–1706, 2011.
21. B. Steinberg. The Černý conjecture for one-cluster automata with prime length cycle. *Theoretical Computer Science*, 412(39):5487–5491, 2011.
22. A. N. Trahtman. An efficient algorithm finds noticeable trends and examples concerning the Černý conjecture. In *MFCS*, volume 4162 of *LNCS*, pp. 789–800. 2006.
23. A. N. Trahtman. Modifying the upper bound on the length of minimal synchronizing word. In *FCT*, volume 6914 of *LNCS*, pp. 173–180. 2011.
24. M. V. Volkov. Synchronizing automata and the Černý conjecture. In *LATA*, volume 5196 of *LNCS*, pp. 11–27. 2008.
25. V. Vorel. Subset Synchronization of Transitive Automata. In *AFL*, pp. 370–381, 2014.