Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Complexity of a problem concerning reset words for Eulerian binary automata [☆]

Vojtěch Vorel

Faculty of Mathematics and Physics, Charles University Malostranské nám. 25, 118 00 Prague, Czech Republic

Article history: Received 12 June 2014 Available online 6 June 2016

ARTICLE INFO

ABSTRACT

A word is called a reset word for a deterministic finite automaton if it maps all the states of the automaton to a unique state. Deciding about the existence of a reset word of a given length for a given automaton is known to be an NP-complete problem. We prove that it remains NP-complete even if restricted to Eulerian automata with binary alphabets as it has been conjectured by Martyugin (2011).

© 2016 Elsevier Inc. All rights reserved.

1. Introduction and preliminaries

A deterministic finite automaton is a triple $A = (Q, X, \delta)$, where Q and X are finite sets and δ is an arbitrary mapping $Q \times X \to Q$. Elements of Q are called *states*, X is the *alphabet*. The *transition function* δ can be naturally extended to $Q \times X^* \to Q$, still denoted by δ . We extend it also by defining

 $\delta(S, w) = \left\{ \delta(s, w) \mid s \in S, w \in X^{\star} \right\}$

for each $S \subseteq Q$. If the automaton is fixed, we write

$$r \xrightarrow{w} s$$

instead of $\delta(r, w) = s$.

For a given automaton $A = (Q, X, \delta)$, we call $w \in X^*$ a reset word if

 $|\delta(Q, w)| = 1.$

If such a word exists, we call the automaton *synchronizing*. Note that each word having a reset word as a factor is also a reset word.

A need for finding reset words appears in several fields of mathematics and engineering. Classical applications (see [14]) include model-based testing, robotic manipulation, and symbolic dynamics, but there are important connections also with information theory [13] and with formal models of biomolecular processes [2].

The Černý Conjecture, a longstanding open problem, claims that each synchronizing automaton has a reset word of length at most $(|Q| - 1)^2$. Though it still remains open, there are many weaker results in this field, see e.g. [11,7] for recent ones.¹

E-mail address: vorel@ktiml.mff.cuni.cz.

http://dx.doi.org/10.1016/j.ic.2016.06.013 0890-5401/© 2016 Elsevier Inc. All rights reserved.

ELSEVIER





 $^{^{}st}$ Research supported by the Czech Science Foundation grant GA14-10799S and the GAUK grant No. 52215.

¹ The result published by Trahtman [12] in 2011 was proved incorrectly, see [6].

Various computational problems arise from the study of synchronization:

- *Given an automaton, decide whether it is synchronizing.* Relatively simple algorithm, which could be traced back to [3], works in polynomial time.
- *Given a synchronizing automaton and a number d, decide whether d is the length of shortest reset words.* This has been shown to be both NP-hard [4] and coNP-hard. More precisely, it is DP-complete [10]. See also [1] and [5] for recent non-approximability results concerning an optimization setting.
- *Given a synchronizing automaton and a number d, decide whether there exists a reset word of length d.* This problem is of our interest. Lying in NP, it is not so computationally hard as the previous problem. However, it is proven to be NP-complete [4]. Following the notation of [9], we call it SYN. Assuming that \mathcal{M} is a class of automata and membership in \mathcal{M} is polynomially decidable, we define a restricted problem:

SYN(\mathcal{M})Input:synchronizing automaton $A \in \mathcal{M}, d \in \mathbb{N}$ Output:does A have a reset word of length d?

An automaton $A = (Q, X, \delta)$ is Eulerian if

$$\sum_{x \in X} |\{r \in Q \mid \delta(r, x) = q\}| = |X|$$

for each $q \in Q$. Informally, there should be exactly |X| transitions incoming to each state. An automaton is *binary* if |X| = 2. The classes of Eulerian and binary automata are denoted by \mathcal{EU} and \mathcal{AL}_2 respectively.

Previous results about various restrictions of SYN can be found in [4,8,9]. Some of these problems turned out to be polynomially solvable, others are NP-complete. In [9] Martyugin conjectured that $SYN(\mathcal{EU} \cap \mathcal{AL}_2)$ is NP-complete. This conjecture is confirmed in the rest of the present paper.

2. Main result

2.1. Proof outline

We prove the NP-completeness of $Syn(\mathcal{EU} \cap \mathcal{AL}_2)$ by a polynomial reduction from 3-SAT. So, for an arbitrary propositional formula ϕ in 3-CNF we construct an Eulerian binary automaton A and a number d such that

 ϕ is satisfiable \Leftrightarrow A has a reset word of length d.

For the rest of the paper we fix a formula

$$\phi = \bigwedge_{i=1}^{m} \bigvee_{\lambda \in C_i} \lambda$$

on *n* variables, where each C_i is a three-element set of literals, i.e., a subset of

$$L_{\phi} = \{x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n\}.$$

We index the literals $\lambda \in L_{\phi}$ by the following mapping κ :

λ	<i>x</i> ₁	<i>x</i> ₂	 x _n	$\neg x_1$	$\neg x_2$	 $\neg x_n$
κ (λ)	0	1	 n — 1	п	n + 1	 2n — 1

Let $A = (Q, X, \delta)$, $X = \{a, b\}$. Because the structure of the automaton A will be very heterogeneous, we use an unusual method of description. The basic principles of the method are:

- We describe the automaton A via a labeled directed multigraph G, representing the automaton in a standard way: edges of G are labeled by single letters a and b and carry the structure of the function δ . Paths in G are thus *labeled* by words from $\{a, b\}^*$.
- There is a collection of labeled directed multigraphs called *templates*. The graph *G* is one of them. Another template is SINGLE, which consists of one vertex and no edges.
- Each template $T \neq SINGLE$ is expressed in a fixed way as a disjoint union through the set $PARTS_T$ of its proper subgraphs (the *parts* of T), extended by a set of additional edges (the *links* of T). Each $H \in PARTS_T$ is isomorphic to some template U. We say that *H* is of type U.

(1)



Fig. 1. Template ABS.



Fig. 2. A barrier of ABS parts.



Fig. 3. Templates CCA, CCI and PIPE(*d*) respectively.

• Let q be a vertex of a template T, lying in a subgraph $H \in \text{PARTS}_{\mathbb{T}}$ which is of type U via a vertex mapping $\rho : H \to U$. The *local address* $\operatorname{adr}_{\mathbb{T}}(q)$ is a finite string of identifiers separated by "]". It is defined inductively by

$$\operatorname{adr}_{\operatorname{T}}(q) = \begin{cases} H \mid \operatorname{adr}_{\operatorname{U}}(\rho(q)) & \text{if } \operatorname{U} \neq \operatorname{SINGLE}, \\ H & \text{if } \operatorname{U} = \operatorname{SINGLE}. \end{cases}$$

The string $adr_G(q)$ is used as a regular vertex identifier.

Having a word $w \in X^*$, we denote a *t*-th letter of *w* by w_t and define the set $S_t = \delta(Q, w_1 \dots w_t)$ of *active states at time t*. Whenever we depict a graph, a solid arrow stands for the label *a* and a dotted arrow stands for the label *b*.

2.2. Description of the graph G

Let us define all the templates and informally comment on their purpose. Fig. 1 defines the template ABS, which does not depend on the formula ϕ .

The state *out* of a part of type ABS is always inactive after application of a word of length at least 2 which does not contain b^2 as a factor. This allows us to ensure the existence of a relatively short reset word. Actually, large areas of the graph (namely the CLAUSE(...) parts) have roughly the shape depicted in Fig. 2, a cylindrical structure with a horizontal barrier of ABS parts. If we use a sufficiently long word with no occurrence of b^2 , the edges outgoing from the ABS parts are never used and almost all states become inactive.

Fig. 3 defines simple templates CCA, CCI and PIPE(d) for a given *length* $d \ge 1$. The activity of an *out* state depends on the last two letters applied. In the case of CCA it is inactive if (and typically only if) the two letters were equal. In the case of CCI it works oppositely, equal letters correspond to active *out* state. One of the key ideas of the entire construction is the following. Let there be a subgraph of the form

part of type PIPE (d)

$$\downarrow a, b$$

part of type CCA or CCI
 $\downarrow a, b$
part of type PIPE (d).
(2)

Before the synchronization process starts, all the states are active. As soon as the second letter of an input word is applied, the activity of the *out* state starts to depend on the last two letters and the pipe below keeps a record of its previous activity. We say that a part H of type PIPE(d) records a sequence $B_1 \dots B_d \in \{0, 1\}^d$ at time t, if it holds that

$$B_k = \mathbf{1} \Leftrightarrow H | s_k \notin S_t.$$



Fig. 4. Templates INC(λ) and NOTINC(λ).



Fig. 5. Template TESTER.

In order to continue with defining templates, let us define a set M_{ϕ} containing all the literals from L_{ϕ} and some auxiliary symbols:

$$M_{\phi} = L_{\phi} \cup \{y_1, \ldots, y_n\} \cup \{z_1, \ldots, z_n\} \cup \{q, q', r, r'\}.$$

We index the 4n + 4 members $\nu \in M_{\phi}$ by the following mapping μ :

ν	q	r	<i>y</i> ₁	<i>x</i> ₁	<i>y</i> ₂	<i>x</i> ₂	 <i>y</i> _n	x _n
$\mu(v)$	1	2	3	4	5	6	2n + 1	2n+2
ν	q'	r'	<i>z</i> ₁	$\neg x_1$	<i>z</i> ₂	$\neg x_2$	 <i>z</i> _n	$\neg x_n$
$\mu(v)$	2n+3	2n + 4	2n + 5	2n + 6	2n + 7	2n + 8	 4n + 3	4n + 4

The inverse mapping is denoted by μ^{-1} . For each $\lambda \in L_{\phi}$ we define templates INC(λ) and NOTINC(λ), both consisting of 12n + 12 SINGLE parts identified by the elements of $\{1, 2, 3\} \times M_{\phi}$. As depicted in Fig. 4a, the links of INC(λ) are:

$(1, \nu) \xrightarrow{a} $	$(2, \lambda)$ $(2, \nu)$	if $v = \lambda$ or $v = r$, otherwise,	$(1, \nu) \xrightarrow{b} \cdot$	(2, r) $(2, v)$	if $v = \lambda$ or $v = r$, otherwise,
$(2, \nu) \xrightarrow{a}$	$(3,q)$ $(3,\nu)$	if $v = r$ or $v = q$, otherwise,	$(2, \nu) \xrightarrow{b} \cdot$	$ \begin{cases} (3,r) \\ (3,\nu) \end{cases} $	if $v = r$ or $v = q$, otherwise.

Note that we use the same identifier for a one-vertex subgraph and for its vertex. As it is clear from Fig. 4b, the links of NOTINC (λ) are:

Fig. 6. Templates FORCER and LIMITER respectively.

$$(1, \nu) \xrightarrow{a} (2, \lambda), \qquad (1, \nu) \xrightarrow{b} (2, r),$$

$$(2, \nu) \xrightarrow{a} \begin{cases} (3, q) & \text{if } \nu = q \text{ or } \nu = \lambda, \\ (3, \nu) & \text{otherwise,} \end{cases} \qquad (2, \nu) \xrightarrow{b} \begin{cases} (3, \lambda) & \text{if } \nu = q \text{ or } \nu = \lambda, \\ (3, \nu) & \text{otherwise.} \end{cases}$$

The key property of such templates comes to light when we need to apply some two-letter word in order to make the state $(3, \lambda)$ inactive assuming (1, r) inactive. If also $(1, \lambda)$ is initially inactive, we can use the word a^2 in both templates. If it is active (which corresponds to the idea of unsatisfied literal λ), we discover the difference between the two templates: The word a^2 works if the type is NOTINC(λ), but fails in the case of INC(λ). Such failure corresponds to the idea of unsatisfied literal λ occurring in a clause of ϕ .

For each *i*-th clause we define the template TESTER (*i*). It consists of 2*n* serially linked parts, namely $level_{\lambda}$ for each $\lambda \in L_{\phi}$, each of type INC(λ) or NOTINC(λ). The particular type of each $level_{\lambda}$ depends on the clause C_i as seen in Fig. 5, so exactly three of them are always of type INC(...). If the corresponding clause is unsatisfied, each of its three literals is unsatisfied, which causes three failures within the levels. Three failures imply at least three occurrences of *b*, which turns up to be too much for a reset word of certain length to exist. Clearly we still need some additional mechanisms to implement this vague vision.

Fig. 6 defines templates FORCER and LIMITER. The idea of template FORCER is simple. Imagine a situation when $q_{1,0}$ or $r_{1,0}$ is active and we need to deactivate the entire forcer by a word of length at most 2n + 3. Any use of *b* would cause an unbearable delay, so if such a word exists, it starts by a^{2n+2} .

The idea of LIMITER is similar, but we tolerate some occurrences of *b* here, namely two of them. This works if we assume $s_{1,0}$ active and it is necessary to deactivate the entire limiter by a word of length at most 6n + 1.

We also need a template PIPES (d, k) for each $d, k \ge 1$. It consists just of k parallel pipes of length d. Namely there is a SINGLE part $s_{d',k'}$ for each d' < d, $k' \le k$ and all the edges are of the form $s_{d',k'} \longrightarrow s_{d'+1,k'}$.

The most complex templates are CLAUSE (*i*) for each $i \in \{1, ..., m\}$. Denote

 $\alpha_i = (i-1)\left(12n-2\right),\,$

 $\beta_i = (m - i) (12n - 2).$

As shown in Fig. 7, CLAUSE (*i*) consists of the following parts:

- Parts sp_1, \ldots, sp_{4n+6} of type SINGLE. The acronym *sp* stands for *spine*.
- Parts *abs*₁,..., *abs*_{4n+6} of type ABS. The entire template has a shape similar to Fig. 2, including the barrier of ABS parts.
 Parts *pipe*_B, *pipe*_F, *pipe*_H of types PIPE(2n 1) and *pipe*_G, *pipe*₁ of types PIPE(2n + 2).
- Parts *cca* and *cci* of types CCA and CCI respectively. Together with the pipes above they realize the idea described in (2). As they form two constellations which work simultaneously, the parts $pipe_G$ and $pipe_1$ typically record mutually inverse sequences. We interpret them as an assignment of the variables x_1, \ldots, x_n . Such assignment is then processed by the tester.
- A part ν of type SINGLE for each $\nu \in M_{\phi}$.
- A part *tester* of type TESTER(*i*).
- A part $\overline{\lambda}$ of type SINGLE for each $\lambda \in L_{\phi}$. While describing the templates INC(λ) and NOTINC(λ) we claimed that in certain case there arises a need to make the state (3, λ) inactive. This happens when the border of inactive area

Fig. 7. Template CLAUSE (*i*).

moves down through the tester levels. The point is that any word of length 6*n* deactivates the entire tester, but we need to ensure that some tester columns, namely the $\kappa(\lambda)$ -th for each $\lambda \in L_{\phi}$, are deactivated one step earlier. If some of them is still active just before the deactivation of tester finishes, the state $\overline{\lambda}$ becomes active, which slows down the synchronization process.

- Parts $pipes_{J}$, $pipes_{K}$ and $pipes_{L}$ of types PIPES (α_{i} , 4n + 4), PIPES (6n 2, 4n + 4) and PIPES (β_{i} , 4n + 4) respectively. There are multiple clauses in ϕ , but multiple testers cannot work in parallel. That is why each of them is padded by a passive PIPES (...) part of size depending on particular *i*. If $\alpha_{i} = 0$ or $\beta_{i} = 0$, the corresponding PIPES part is not present in CLAUSE (*i*).
- Parts pipe_A, pipe_C, pipe_D, pipe_E of types PIPE(12mn + 4n − 2m + 6), PIPE(4), PIPE(α_i + 6n − 1), PIPE(β_i) respectively.
- A part *forcer* of type FORCER. This part guarantees that only the letter *a* is used in certain segment of the word *w*. This is necessary for the data produced by *cca* and *cci* to safely leave the parts $pipe_F$, $pipe_H$ and line up in the states of the form ν for $\nu \in M_{\phi}$, from where they are shifted to the tester.
- A part *limiter* of type LIMITER. This part guarantees that the letter *b* occurs at most twice when the border of inactive area passes through the tester. Because each unsatisfied literal from the clause requests an occurrence of *b*, only a satisfied clause meets all the conditions for a reset word of certain length to exist.

See also Table 1. Links of CLAUSE(i) that may not be clear from Fig. 7 are

Fig. 8. The graph G.

$$\nu \xrightarrow{a} \begin{cases} pipes_{J}|s_{1,\mu(\nu)} & \text{if } \nu \in \{x_{n}, \neg x_{n}\}, \\ \mu^{-1}(\mu(\nu)+1) & \text{otherwise}, \end{cases} \quad \nu \xrightarrow{b} pipes_{J}|s_{1,\mu(\nu)} \end{cases}$$

for each $v \in M_{\phi}$ and

$$pipes_{L}|s_{\beta_{i},k} \xrightarrow{a,b} \begin{cases} \overline{\mu^{-1}(k)} & \text{if } \mu^{-1}(k) \in L_{\phi}, \\ abs_{k+2}|in & \text{otherwise}, \end{cases} \xrightarrow{\overline{\lambda} \xrightarrow{a,b}} abs_{\mu(\lambda)+2}|in$$

for each $k \in \{1, ..., 4n + 4\}$, $\lambda \in L_{\phi}$. Note that we rely on the addressing system introduced in Section 2.1.

We are ready to form the whole graph *G*, see Fig. 8. For each $i \in \{1, ..., m\}$ there are parts cl_i , abs_i of types CLAUSE (i) and ABS, respectively, and parts q_i , r_i , r'_i of type SINGLE. The parts s_1 , s_2 are of type SINGLE as well. The edge incoming to a cl_i part ends in $cl_i|sp_1$, the outgoing *b*-edge one starts in $cl_i|sp_{4n+6}$. When no states outside ABS parts are active within each CLAUSE (...) part and no *out*, r_1 nor r_2 state is active in any ABS part, the word b^2ab^{4n+m+7} takes all active states to s_2 and completes the synchronization.

Graph G does not fully represent the automaton A yet because there are

- 8mn + 4m vertices with only one outgoing edge, namely $cl_i | abs_k | out$ and $cl_i | sp_l$ for each $i \in \{1, ..., m\}$, $k \in \{1, ..., 4n + 6\}$, $l \in \{7, ..., 4n + 4\}$,
- 8mn + 4m vertices with only one incoming edge: $cl_i | v$ and $cl_i | pipes_J | (1, v')$ for each $i \in \{1, ..., m\}$, $v \in M_{\phi} \setminus \{q, q'\}$, $v' \in M_{\phi} \setminus \{x_n, \neg x_n\}$.

We do not need to specify the missing edges exactly, let us just say that they somehow connect the relevant states so that the automaton A is complete and Eulerian. Let us set

d = 12mn + 8n - m + 18

and prove that the equivalence (1) holds.

2.3. From an assignment to a word

First, let us suppose that there is an assignment $\xi_1, \ldots, \xi_n \in \{0, 1\}$ of the variables x_1, \ldots, x_n (respectively) satisfying the formula ϕ and prove that the automaton A has a reset word w of length d. For each $j \in \{1, \ldots, n\}$ we denote

$$\sigma_j = \begin{cases} a & \text{if } \xi_j = \mathbf{1}, \\ b & \text{if } \xi_j = \mathbf{0} \end{cases}$$

and for each $i \in \{1, ..., m\}$ we choose a satisfied literal $\overline{\lambda}_i$ from C_i . We set

$$w = a^{2} (\sigma_{n}a) (\sigma_{n-1}a) \dots (\sigma_{1}a) aba^{2n+3} b \left(a^{6n-2} v_{1} \right) \dots \left(a^{6n-2} v_{m} \right) b^{2} a b^{4n+m+7},$$

where for each $i \in \{1, ..., m\}$ we use the word

$$v_i = u_{i,x_1} \dots u_{i,x_n} u_{i,\neg x_1} \dots u_{i,\neg x_n},$$

denoting

$$u_{i,\lambda} = \begin{cases} a^3 & \text{if } \lambda = \overline{\lambda}_i \text{ or } \lambda \notin C_i, \\ ba^2 & \text{if } \lambda \neq \overline{\lambda}_i \text{ and } \lambda \in C_i \end{cases}$$

for each $\lambda \in L_{\phi}$. We see that $|v_i| = 6n$ and therefore

$$|w| = 4n + 8 + m(12n - 2) + 4n + m + 10 = 12mn + 8n - m + 18 = d.$$

Let us denote

$$\gamma = 12mn + 4n - 2m + 9$$

and

$$\overline{S}_t = Q \setminus S_t$$

for each $t \leq d$. Because the first occurrence of b^2 in *w* starts at the γ -th position, we have:

Lemma 2.1. Each state of the form $cl_{\dots}|abs_{\dots}|out$ or $abs_{\dots}|out$ lies in $\overline{S}_2 \cap \cdots \cap \overline{S}_{\gamma}$.

Let us fix an arbitrary $i \in \{1, ..., m\}$ and describe a growing area of inactive states within cl_i . We use the following method of verifying inactivity of states: Having a state $s \in Q$ and $t, k \ge 1$ such that any path of length k ending in s contains a member of $\overline{S}_{t-k} \cap \cdots \cap \overline{S}_{t-1}$, we easily deduce that $s \in \overline{S}_t$. In such case let us just say that k witnesses that $s \in \overline{S}_t$. The following claims follow directly from the definition of w. Note that Claim 7 relies on the fact that b occurs only twice in v_i :

Lemma 2.2.

- 1. $\{cl_i|sp_1,\ldots,cl_i|sp_{4n+6}\} \subseteq \overline{S}_2 \cap \cdots \cap \overline{S}_{\gamma}$
- 2. $cl_i | pipe_B \cup cl_i | pipe_F \cup cl_i | pipe_H \subseteq \overline{S}_{2n+1} \cap \cdots \cap \overline{S}_{\gamma}$
- 3. $cl_i | cca \cup cl_i | cci \cup cl_i | pipe_C \subseteq \overline{S}_{2n+5} \cap \cdots \cap \overline{S}_{\gamma}$
- 4. $cl_i | pipe_G \cup cl_i | pipe_I \cup cl_i | forcer \subseteq \overline{S}_{4n+7} \cap \cdots \cap \overline{S}_{\gamma}$
- 5. $\{cl_i | \nu : \nu \in M_{\phi}\} \subseteq \overline{S}_{4n+8} \cap \cdots \cap \overline{S}_{\gamma}$
- 6. $cl_i | pipes_{I} \cup cl_i | pipes_{K} \cup cl_i | pipe_{D} \subseteq \overline{S}_{10n+\alpha_i+6} \cap \cdots \cap \overline{S}_{\gamma}$
- 7. $cl_i | limiter \cup cl_i | tester$ $\subseteq \overline{S}_{16n+\alpha_i+6} \cap \cdots \cap \overline{S}_{\gamma}$
- 8. $cl_i | pipe_A \cup cl_i | pipe_E \cup cl_i | pipes_L \subseteq \overline{S}_{\gamma-1} \cap \overline{S}_{\gamma}$

Proof. 1. Claim: $\{cl_i | sp_1, \ldots, cl_i | sp_{4n+6}\} \subseteq \overline{S}_2 \cap \cdots \cap \overline{S}_{\gamma}$.

We have $w_1 w_2 = a^2$ and there is no path labeled by a^2 ending in any $cl_i | sp_{...}$ state, so such states lie in \overline{S}_2 . For each $t = 3, ..., \gamma$ we can inductively use k = 1 to witness the memberships in \overline{S}_t . In the induction step we use Lemma 2.1, which excludes the *out* states of the ABS parts from each corresponding \overline{S}_{t-1} .

- 2. Claim: $cl_i | pipe_B \cup cl_i | pipe_F \cup cl_i | pipe_H \subseteq \overline{S}_{2n+1} \cap \cdots \cap \overline{S}_{\gamma}$. All the memberships are witnessed by k = 2n - 1, because any path of the length 2n - 1 ending in such state must use a $cl_i | sp_{\dots}$ state and such states lie in $\overline{S}_2 \cap \cdots \cap \overline{S}_{\gamma}$ by the previous claim.
- 3. Claim: cl_i|cca ∪ cl_i|cci ∪ cl_i|pipe_C ⊆ S_{2n+5} ∩ ··· ∩ S_γ. We have w_{2n+2}... w_{2n+5} = a²ba, which clearly maps each state of cl_i|cca, cl_i|cci or cl_i|pipe_C out of those parts. Each path of length 4 leading into the parts from outside starts in S_{2n+1}, so it follows that all the states lie in S_{2n+5}. To prove the rest we inductively use the witness k = 1.
 4. Claim: cl_i|pipe_G ∪ cl_i|pipe_I ∪ cl_i|forcer ⊆ S_{4n+7} ∩ ··· ∩ S_γ.
- In the cases of $cl_i|pipe_G \circ cl_i|pipe_I \circ cl_i|pipe_I$ we just use the witness k = 2n + 2. In the case of $cl_i|forcer$ we proceed in the same way as in the previous claim. Because $w_{2n+5} \dots w_{4n+7} = a^{2n+3}$, only the states $q_{\dots,0}$ can be active within the part $cl_i|forcer$ in time 2n + 6. The word $w_{2n+7} \dots w_{4n+7}$ maps all such states out of $cl_i|forcer$. Each path of length 2n + 2

505

leading into cl_i [forcer from outside starts in \overline{S}_{2n+5} , so it follows that all states from cl_i [forcer lie in \overline{S}_{4n+7} . To handle $t = 4n + 8, ..., \gamma$ we inductively use the witness k = 1.

5. Claim: $\{cl_i | v : v \in M_{\phi}\} \subseteq \overline{S}_{4n+8} \cap \cdots \cap \overline{S}_{\gamma}$. In the cases of $cl_i | q$ and $cl_i | q'$ we use the witness 1. We have $w_{4n+8} = b$ and the only edges labeled by b incoming to remaining states could be some of the 8mn + 4m unspecified edges of G. But we have $w_{4n+6}w_{4n+7} = a^2$, so each out state of any ABS part lies in \overline{S}_{4n+7} and thus no unspecified edge starts in a state outside \overline{S}_{4n+7} .

- 6. Claim: $cl_i|pipes_J \cup cl_i|pipes_K \cup cl_i|pipe_D \subseteq \overline{S}_{10n+\alpha_i+6} \cap \cdots \cap \overline{S}_{\gamma}$. We use witnesses $k = \alpha_i$ for $cl_i|pipes_J$, k = 6n - 2 for $cl_i|pipes_K$ and $k = \alpha_i + 6n - 1$ for $cl_i|pipe_D$.
- 7. *Claim:* $cl_i | limiter \cup cl_i | tester \subseteq \overline{S}_{16n+\alpha_i+6} \cap \cdots \cap \overline{S}_{\gamma}$. Because

$$w_{4n+\alpha_i+9}\ldots w_{10n+\alpha_i+6}=a^{6n-2},$$

there are only states of the form $cl_i|limiter|s_{...,0}$ in the intersection of $cl_i|limiter$ and $S_{10n+\alpha_i+6}$. Together with the fact that there are only two occurrences of *b* in v_i it confirms that the case of $cl_i|limiter$ holds. The case of $cl_i|tester$ is easily witnessed by k = 6n.

8. Claim: $cl_i | pipe_A \cup cl_i | pipe_E \cup cl_i | pipe_L \subseteq \overline{S}_{\gamma-1} \cap \overline{S}_{\gamma}$. We use witnesses k = 12mn + 4n - 2m + 6 for $cl_i | pipe_A$ and $k = \beta_i$ for $cl_i | pipe_E$, $cl_i | pipe_L$. \Box

For each $\lambda \in L_{\phi}$ we ensure by the word $u_{i,\lambda}$ that the $\kappa(\lambda)$ -th tester column is deactivated in advance, namely at time $t = 16n + \alpha_i + 5$. The advance allows the following key claim to hold true:

Lemma 2.3. $\{cl_i | \overline{\lambda} : \lambda \in L_{\phi}\} \subseteq \overline{S}_{\gamma-1} \cap \overline{S}_{\gamma}.$

Proof. For each such λ we choose

$$k = 6n - 3\kappa(\lambda) + \beta_i + 1$$

as a witness of $cl_i|\overline{\lambda} \in \overline{S}_{\gamma-1}$. There is only one state where a path of length k ending in $\overline{\lambda}$ starts: the state

 $s = cl_i |tester| |level_{\lambda}| (3, \lambda)$.

It holds that

$$s \in \overline{S}_{10n+\alpha_i+3\kappa(\lambda)+6} \cap \cdots \cap \overline{S}_{\gamma},$$

as is easily witnessed by $k' = 3\kappa(\lambda)$ using Claim 6 of Lemma 2.2. But we are going to show also that

$$s \in \overline{S}_{10n+\alpha_i+3\kappa(\lambda)+5},$$

which will imply that k is a true witness of $\overline{\lambda} \in \overline{S}_{\gamma-1}$, because

 $(\gamma - 1) - k = 10n + \alpha_i + 3\kappa(\lambda) + 5.$

So let us prove the membership (3). We need to observe, using the definition of w, that:

- At time 2n + 5 the part $pipe_{G}$ records the sequence
 - $0, 1, \xi_1, \xi_1, \xi_2, \xi_2, \ldots, \xi_n, \xi_n$

and the part pipe_I records the sequence of inverted values. Because

$$w_{2n+6}\ldots w_{4n+7}=a^{2n+2},$$

at time 4n + 7 the states q, r' are active, the states q', r are inactive and for each $j \in \{1, ..., n\}$ it holds that

$$x_j \in \overline{S}_{4n+7} \Leftrightarrow y_j \in \overline{S}_{4n+7} \Leftrightarrow \neg x_j \in S_{4n+7} \Leftrightarrow z_j \in S_{4n+7} \Leftrightarrow \xi_j = \mathbf{1}.$$

Because $w_{4n+8} = b$, at time $10n + \alpha_i + 6$ we find the whole structure above shifted to the first row of $cl_i | tester$, so particularly for $\lambda \in L_{\phi}$:

 $cl_i|tester|level_{x_1}|(1,\lambda) \in \overline{S}_{10n+\alpha_i+6} \Leftrightarrow \lambda \text{ is satisfied by } \xi_1, \dots, \xi_n.$

• From a simple induction on tester levels it follows that

 $cl_i|tester|level_{\lambda}|(1,r) \in \overline{S}_{10n+\alpha_i+3\kappa(\lambda)+3}.$

(3)

Note that

 $w_{10n+\alpha_i+3\kappa(\lambda)+4}w_{10n+\alpha_i+3\kappa(\lambda)+5}w_{10n+\alpha_i+3\kappa(\lambda)+6} = u_{i,\lambda}$

and distinguish the following cases:

• If $\lambda = \overline{\lambda}_i$, we have $\lambda \in C_i$, the part $cl_i | tester | level_{\lambda}$ is of type INC (λ) and $u_{i,\lambda} = a^3$. We also know that λ is satisfied, so

 $cl_i|tester|level_{x_1}|(1, \lambda) \in \overline{S}_{10n+\alpha_i+6}.$

The state above is the only state, from which any path of length $3\kappa(\lambda) - 3$ leads to $cl_i|tester|level_{\lambda}|(1, \lambda)$, so we deduce that

 $cl_i|tester|level_{\lambda}|(1,\lambda) \in \overline{S}_{10n+\alpha_i+3\kappa(\lambda)+3}.$

We see that each path labeled by a^2 ending in $cl_i|tester|level_{\lambda}|(3,\lambda)$ starts in $cl_i|tester|level_{\lambda}|(1,\lambda)$ or in $cl_i|tester|level_{\lambda}|(1,r)$, but each of the two states lies in $\overline{S}_{10n+\alpha_i+3\kappa(\lambda)+3}$. So the membership (3) holds.

- If $\lambda \notin C$, the part $cl_i|tester|level_{\lambda}$ is of type NOTINC (λ) and $u_{i,\lambda} = a^3$. Particularly $w_{10n+\alpha_i+3\kappa(\lambda)+5} = a$ but no edge labeled by a comes to $cl_i|tester|level_{\lambda}|(3, \lambda)$ and the membership (3) follows trivially.
- If $\lambda \neq \overline{\lambda_i}$ and $\lambda \in C_i$, the part $cl_i | tester | level_{\lambda}$ is of type INC (λ) and $u_{i,\lambda} = ba^2$. Particularly

 $w_{10n+\alpha_i+3\kappa(\lambda)+4}w_{10n+\alpha_i+3\kappa(\lambda)+5} = ba,$

but no path labeled by ba comes to $cl_i |tester| level_{\lambda}|$ (3, λ), so we reach the same conclusion as in the previous case.

We have proven that $cl_i|_{\overline{\lambda}}$ lies in $\overline{S}_{\gamma-1}$. From Claim 8 of Lemma 2.2 it follows directly that it lies also in \overline{S}_{γ} .

We see that within cl_i only the states from the ABS parts can lie in $S_{\gamma-1}$. Since $w_{\gamma-2}w_{\gamma-1} = a^2$, no state r_1 , r_2 nor out from any ABS part lies in $S_{\gamma-1}$. Now we easily check that all the states possibly present in $S_{\gamma-1}$ are mapped to s_2 by the word $w_{\gamma} \dots w_d = b^2 a b^{4n+m+7}$.

2.4. From a word to an assignment

From now on we suppose that there is a reset word w of length

d = 12mn + 8n - m + 18.

The following lemma is not hard to verify:

Lemma 2.4.

- 1. Up to labeling there is a unique pair of paths, both of a length $l \le d 2$, leading from $cl_1|pipe_A|s_1$ and $cl_2|pipe_A|s_1$ to a common end. They are of length d 2 and meet in s_2 .
- 2. The word w starts by a^2 .
- **Proof.** 1. The leading segments of both paths are similar since they stay within the parts cl_1 and cl_2 :

$$pipe_{A}|s_{1} \xrightarrow{a,b} \dots \xrightarrow{a,b} pipe_{A}|s_{12mn+4n-2m+6} \xrightarrow{a,b} abs_{1}|in \xrightarrow{b} \dots \xrightarrow{b} abs_{1}|r_{1} \xrightarrow{b} abs_{1}|out \xrightarrow{a} sp_{1} \xrightarrow{b} \dots \xrightarrow{b} sp_{4n+6}.$$

Once the paths leave the parts cl_1 and cl_2 , the shortest way to merge is the following:

$$cl_{1}|sp_{4n+6} \xrightarrow{b} q_{1} \xrightarrow{b} q_{2} \xrightarrow{b} \dots \xrightarrow{b} q_{m-1} \xrightarrow{b} q_{m} \xrightarrow{b} s_{1} \xrightarrow{b} s_{2}$$
$$cl_{2}|sp_{4n+6} \xrightarrow{b} q_{2} \xrightarrow{b} q_{3} \xrightarrow{b} \dots \xrightarrow{b} q_{m} \xrightarrow{b} s_{1} \xrightarrow{b} s_{2} \xrightarrow{b} s_{2}$$

Having the description above it is easy to verify that the length is d - 2 and there is no way to make the paths shorter. 2. Suppose that $w_1w_2 \neq a^2$. Any of the three possible values of w_1w_2 implies that

$$\{cl_i|sp_3,\ldots,cl_i|sp_{4n+6}\}\subseteq S_2$$

for each *i*. It cannot hold that $w = w_1 w_2 b^{d-2}$, because in such case all $cl_{...}|cca|s_b$ states would be active in any time $t \ge 3$. So the word *w* has a prefix $w_1 w_2 b^k a$ for some $k \ge 0$. If $k \le 4n + 3$, it holds that $cl_i|sp_{4n+6} \in S_{k+2}$ and therefore

$$d - (4n + 4) \ge d - k$$

leading from such two states to a common end. \Box

The second claim implies that $cl_i | pipe_A | s_1 \in S_2$ for each $i \in \{1, ..., m\}$, so it follows that

 $\delta(Q, w) = \{s_2\}.$

Let us denote

$$\overline{d} = 12mn + 4n - 2m + 11,$$

$$w = w_1 \dots w_{\overline{d}}$$

and

$$S' = \left\{ s \in Q \mid (\exists d \in \mathbb{N}) \, \delta\left(s, b^d\right) = s_2 \right\}.$$

The following lemma holds because no edges labeled by a are available for final segments of the paths described in the first claim of Lemma 2.4:

Lemma 2.5.

- 1. The word w can be written as $w = \overline{w}b^{4n+m+7}$ for some word \overline{w} .
- 2. For any $t \ge \overline{d}$, no state from any cl_{\dots} part lies in S_t , except for the sp_ states.

Proof. 1. Let us write $w = w_1 w_2 w'$. From Lemma 2.4 it follows that

 $\delta\left(cl_1|pipe_A|s_1, w'\right) = \delta\left(cl_2|pipe_A|s_1, w'\right)$

and w' has to label some of the paths determined up to the labeling in Lemma 2.4(1). The final 4n + m + 7 edges of the paths lead from $cl_1|sp_1$ and $cl_2|sp_1$ to s_2 . All the transitions used here are necessarily labeled by b.

2. The claim is easy to observe, since the first claim implies that S_t is a subset of S'. \Box

The next lemma is based on properties of the parts $cl_{...}$ [forcer but to prove that no more *a* follows the enforced factor a^{2n+1} we also need to observe that each $cl_{...}$ [cca|out or each $cl_{...}$]cci|out lies in S_{2n+4} :

Lemma 2.6. The word \overline{w} starts by $\overline{u}a^{2n+1}b$ for some \overline{u} of length 2n + 6.

Proof. At first we prove that \overline{w} starts by $\overline{u}a^{2n+1}$. Lemma 2.4(2) implies that $cl_1|pipe_B|s_1 \in S_2$, so obviously at least one of the states $cl_1|forcer|q_{1,0}$ and $cl_1|forcer|r_{1,0}$ lies in S_{2n+6} . If $w_{2n+6+k} = b$ for some $k \in \{1, ..., 2n+1\}$, it holds that $cl_i|forcer|q_{k,2}$ or $cl_i|forcer|r_{k,2}$ lies in S_{2n+6+k} . From such a state no path of length at most 2n + 3 - k leads to $cl_i|pipe_D|s_1$ and therefore no path of length at most

$$(2n+3-k) + (\alpha_i + 6n - 1) + (6n - 2) + \beta_i + 3 = \overline{d} - (2n+6+k)$$

leads into S', which contradicts Lemma 2.5(2). It remains to show that there is b after the prefix $\overline{u}a^{2n+1}$. Lemma 2.4(2) implies that both $cl_1|cca|in$ and $cl_1|cci|in$ lie in S_{2n+1} , from which it is not hard to deduce that $cl_1|cca|out$ or $cl_1|cci|out$ lies in S_{2n+4} and therefore $cl_1|q$ or $cl_1|r$ lies in S_{4n+7} . Any path of length $\overline{d} - (4n+7)$ leading from $cl_1|q$ or $cl_1|r$ into \overline{S} starts by an edge labeled by b. \Box

Now we are able to write the word \overline{w} as

$$\overline{w} = \overline{u}a^{2n+1}b\left(\overline{v}_1v_1'c_1\right)\ldots\left(\overline{v}_mv_m'c_m\right)w_{\overline{d}-2}w_{\overline{d}-1}w_{\overline{d}},$$

where $|\overline{v}_k| = 6n-2$, $|v'_k| = 6n-1$ and $|c_k| = 1$ for each k and denote $d_i = 10n + \alpha_i + 6$. At time 2n+5 the parts $cl_{...}|pipe_G$ and $cl_{...}|pipe_I$ record mutually inverse sequences. The factor $a^{2n+1}b$ after \overline{u} guarantees that at time d_i we find the information pushed to the first rows of testers:

Lemma 2.7. For each $i \in \{1, ..., m\}$, $j \in \{1, ..., n\}$ it holds that

$$cl_{i}|tester|level_{x_{1}}|(1, x_{j}) \in S_{d_{i}} \Leftrightarrow$$
$$cl_{i}|tester|level_{x_{1}}|(1, \neg x_{j}) \notin S_{d_{i}} \Leftrightarrow w_{2n-2j+2} \neq w_{2n-2j+3}$$

Proof. From the definition of CCA and CCI it follows that at time 2n + 5 the parts $pipe_G$ and $pipe_I$ record the sequences $B_{(2n+3)} \dots B_{(2)}$ and $B'_{(2n+3)} \dots B'_{(2)}$ respectively, where

$$B_{(k)} = \begin{cases} 1 & \text{if } w_k = w_{k+1}, \\ 0 & \text{otherwise}, \end{cases} \qquad B'_{(k)} = \begin{cases} 0 & \text{if } w_k = w_{k+1}, \\ 1 & \text{otherwise}. \end{cases}$$

Whatever the letter w_{2n+6} is, Lemma 2.6 implies that

$$cl_i | x_i \in S_{4n+7} \Leftrightarrow cl_i | \neg x_i \notin S_{4n+7} \Leftrightarrow w_{2n-2i+2} \neq w_{2n-2i+3},$$

from which the claim follows easily according to the shape of $cl_i|pipes_I$ and $cl_i|pipes_K$. \Box

Let us define the assignment $\xi_1, \ldots, \xi_n \in \{0, 1\}$:

$$\xi_j = \begin{cases} 1 & \text{if } w_{2n-2j+2} = w_{2n-2j+3}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the definition may be directly rephrased according to Lemma 2.7. The following lemma holds due to the *cl*...|*limiter* parts:

Lemma 2.8. For each $i \in \{1, ..., m\}$ there are at most two occurrences of b in the word v'_i .

Proof. It is easy to see that $cl_i | limiter | s_{1,0} \in S_{10n+\alpha_i+6}$ and to note that

$$v'_i = w_{10n+\alpha_i+7} \dots w_{16n+\alpha_i+5}.$$

Within the part cl_i limiter no state except for $s_{6n-2,0}$ can lie in $S_{16n+\alpha_i+5}$, because from such states there is no path of length at most

$$\overline{d} - (16n + \alpha_i + 5) = \beta_i + 4$$

leading into S'.

The shortest paths from $s_{1,0}$ to $s_{6n-2,0}$ have length 6n-3 and each path from $s_{1,0}$ into S' uses the state $s_{6n-2,0}$. So there is a path P leading from $s_{1,0}$ to $s_{6n-2,0}$ labeled by a prefix of v'. We distinguish the following cases:

- If *P* is of length 6n 3, we just note that such path is unique and labeled by a^{6n-3} . No *b* occurs in v' except for the last two positions.
- If *P* is of length 6n 2, it uses an edge of the form $s_{k,0} \xrightarrow{b} s_{k+1,1}$. Such edges preserve the distance to s_{6n-2} , so the rest of *P* must be a shortest path from $s_{k+1,1}$ to $s_{6n-2,0}$. Such paths are unique and labeled by a^{6n-2-k} . Any other *b* can occur only at the last position.
- If *P* is of length 6n 1, it is labeled by whole v'. Because any edge labeled by *b* preserves or increases the distance to s_{6n-2} , the path *P* can use at most two of them. \Box

Now we choose any $i \in \{1, ..., m\}$ and prove that the assignment $\xi_1, ..., \xi_n$ satisfies the clause $\bigvee_{\lambda \in C_i} \lambda$. Let $p \in \{0, 1, 2, 3\}$ denote the number of unsatisfied literals in C_i .

As we claimed before, all tester columns corresponding to any $\lambda \in L_{\phi}$ have to be deactivated earlier than other columns. Namely, if $cl_i|tester|level_{x_1}|(1,\lambda)$ is active at time d_i , which happens if and only if λ is not satisfied by ξ_1, \ldots, ξ_n , the word v'_ic_i must not map it to $cl_i|pipes_L|s_{1,\mu(\lambda)}$. If $cl_i|tester|level_{\lambda}$ is of type INC (λ), the only way to ensure this is to use the letter b when the border of inactive area lies at the first row of $cl_i|tester|level_{\lambda}$. Thus each unsatisfied $\lambda \in C_i$ implies an occurrence of b in corresponding segment of v'_i :

Lemma 2.9. There are at least p occurrences of the letter b in the word v'_i .

Proof. Let $\lambda_1, \ldots, \lambda_p$ be the unsatisfied literals of C_i . From Lemma 2.7 it follows easily that

$$cl_i|tester|level_{\lambda_k}|(1,\lambda_k) \in S_{d_i+3\kappa(\lambda_k)}$$

for each $k \in \{1, ..., p\}$. The part $cl_i | tester | level_{\lambda_k}$ is of type INC (λ_k), which implies that any path of the length

$$\left(\overline{d}-3\right)-\left(d_{i}+3\kappa\left(\lambda_{k}\right)\right)$$

starting by *a* takes $cl_i|tester|level_{\lambda_k}|(1,\lambda_k)$ to the state $cl_i|\overline{\lambda}$, which lies outside $S_{\overline{d}-3}$, as it is implied by Lemma 2.5(2). We deduce that $w_{d_i+3\kappa(\lambda_k)+1} = b$. \Box

By Lemma 2.8 there are at most two occurrences of b in v'_i , so we get $p \le 2$ and there is at least one satisfied literal in C_i .

References

- M.V. Berlinkov, On two algorithmic problems about synchronizing automata, in: A. Shur, M. Volkov (Eds.), Developments in Language Theory, in: Lecture Notes in Computer Science, vol. 8633, Springer International Publishing, 2014, pp. 61–67.
- [2] P. Bonizzoni, N. Jonoska, Existence of constants in regular splicing languages, Inf. Comput. 242 (2015) 340-353.
- [3] J. Černý, Poznámka k homogénnym experimentom s konečnými automatmi, Mat.-Fyz. Čas. 14 (3) (1964) 208-216.
- [4] D. Eppstein, Reset sequences for monotonic automata, SIAM J. Comput. 19 (3) (1990) 500-510.
- [5] P. Gawrychowski, D. Straszak, Strong inapproximability of the shortest reset word, in: G. Italiano, G. Pighizzini, D. Sannella (Eds.), Mathematical Foundations of Computer Science 2015, in: Lecture Notes in Computer Science, vol. 9234, Springer, Berlin, Heidelberg, 2015, pp. 243–255.
- [6] F. Gonze, R.M. Jungers, A. Trakhtman, A note on a recent attempt to improve the Pin-Frankl bound, Discret. Math. Theor. Comput. Sci. 17 (1) (2015) 307–308.
- [7] M. Grech, A. Kisielewicz, The Černý conjecture for automata respecting intervals of a directed graph, Discret. Math. Theor. Comput. Sci. 15 (3) (2013) 61–72.
- [8] P.V. Martugin, Complexity of problems concerning reset words for some partial cases of automata, Acta Cybern. 19 (2) (2009) 517–536.
- [9] P. Martyugin, Complexity of problems concerning reset words for cyclic and eulerian automata, Theor. Comput. Sci. 450 (2012) 3–9.
- [10] J. Olschewski, M. Ummels, The complexity of finding reset words in finite automata, in: P. Hliněný, A. Kučera (Eds.), Mathematical Foundations of Computer Science 2010, in: Lecture Notes in Computer Science, vol. 6281, Springer, Berlin, Heidelberg, 2010, pp. 568–579.
- [11] B. Steinberg, The Černý conjecture for one-cluster automata with prime length cycle, Theor. Comput. Sci. 412 (39) (2011) 5487–5491.
- [12] A. 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, in: Lecture Notes in Computer Science, vol. 6914, Springer, Berlin, Heidelberg, 2011, pp. 173–180.
- [13] N. Travers, J. Crutchfield, Exact synchronization for finite-state sources, J. Stat. Phys. 145 (5) (2011) 1181–1201.
- [14] M.V. Volkov, Synchronizing automata and the Černý conjecture, in: C. Martín-Vide, F. Otto, H. Fernau (Eds.), Language and Automata Theory and Applications, in: Lecture Notes in Computer Science, vol. 5196, Springer, Berlin, Heidelberg, 2008, pp. 11–27.