Finitely generated ideal languages and synchronizing automata

Vladimir V. Gusev, Marina I. Maslennikova, Elena V. Pribavkina

Ural Federal University, Ekaterinburg, Russia
vl.gusev@gmail.com, maslennikova.marina@gmail.com,
elena.pribavkina@usu.ru

Abstract. We study representations of ideal languages by means of strongly connected synchronizing automata. For every finitely generated ideal language L we construct such an automaton with at most 2^n states, where n is the maximal length of words in L. Our constructions are based on the De Bruijn graph.
Keywords: ideal language, synchronizing automaton, synchronizing word, reset complexity.

1 Introduction

Let $\mathscr{A} = \langle Q, \Sigma, \delta\rangle$ be a *deterministic finite automaton* (DFA for short), where Q is the *state set*, Σ stands for the *input alphabet*, and $\delta : Q \times \Sigma \to Q$ is the *transition function* defining an action of the letters in Σ on Q. When δ is clear from the context, we will write $q\,.\,w$ instead of $\delta(q, w)$ for $q \in Q$ and $w \in \Sigma^*$.

A DFA $\mathscr{A} = \langle Q, \Sigma, \delta\rangle$ is called *synchronizing* if there exists a word $w \in \Sigma^*$ which leaves the automaton in unique state no matter at which state in Q it is applied: $q\,.\,w = q'\,.\,w$ for all $q, q' \in Q$. Any word w with such property is said to be *synchronizing* (or *reset*) word for the DFA $\mathscr{A}$. For the last 50 years synchronizing automata received a great deal of attention. For a brief introduction to the theory of synchronizing automata we refer the reader to the recent surveys [7, 8].

In the present paper we focus on language theoretic aspects of the theory of synchronizing automata. We denote by $\mathrm{Syn}(\mathscr{A})$ the language of synchronizing words for a given automaton $\mathscr{A}$. It is well known that $\mathrm{Syn}(\mathscr{A})$ is regular [8]. Furthermore, it is an *ideal* in Σ^*, i.e. $\mathrm{Syn}(\mathscr{A}) = \Sigma^* \mathrm{Syn}(\mathscr{A})\Sigma^*$. On the other hand, every ideal language L serves as a language of synchronizing words for some automaton. For instance, the minimal automaton of the language L is synchronized by L [4]. Thus, synchronizing automata can be considered as a special representation of ideal languages. Effectiveness of such representation was addressed in [4]. The *reset complexity* $rc(L)$ of an ideal language L is the minimal possible number of states in a synchronizing automaton $\mathscr{A}$ such that $\mathrm{Syn}(\mathscr{A}) = L$. Every such automaton $\mathscr{A}$ is called *minimal synchronizing automaton* (for brevity, MSA). Let $sc(L)$ be the number of states in the minimal automaton recognizing L. Then for every ideal language L we have $rc(L) \le sc(L)^\star$. Moreover, there are

* since the minimal automaton is synchronized by L

languages L_n for every $n \geq 3$ such that $rc(L_n) = n$ and $sc(L_n) = 2^n - n$, see [4]. Thus, representation of an ideal language by means of a synchronizing automaton can be exponentially smaller than “traditional” representation via minimal automaton. However, no reasonable algorithm is known for computing MSA of a given language. One of the obstacles is that MSA is not uniquely defined. For instance, there is a language with at least two different MSA's: one of them is strongly connected, another one has a sink state [4]. Therefore, some refinement of the notion of MSA seems to be necessary. Another important observation is the following: minimal synchronizing automata for the aforementioned languages L_n are strongly connected. Thus, one may expect that there is always a strongly connected MSA for an ideal language. In the present paper we show that it is not the case. Moreover, the smallest strongly connected automaton with a language L as the language of synchronizing words may be exponentially larger than a minimal synchronizing automaton of L.

Another source of motivation for studying representations of ideal languages by means of synchronizing automata comes from the famous Černý conjecture. Černý already in 1964 conjectured that every synchronizing automaton possesses a synchronizing word of length at most $(n-1)^2$. Despite intensive efforts of researchers this conjecture is still widely open. We can restate the Černý conjecture in terms of reset complexity as follows: if ℓ is the minimal length of words in an ideal language L then $rc(L) \geq \sqrt{\ell} + 1$. Thus, we hope that deeper understanding of reset complexity will bring us new ideas to resolve this long standing conjecture. It is well known that the Černý conjecture holds true whenever it holds true for strongly connected automata. In this regard an interesting related question was posed in [2]: does every ideal language serve as the language of synchronizing words for some strongly connected automaton? For instance, if the answer is negative then there is a way to simplify formal language statement of the Černý conjecture. Unfortunately, it is not the case. Recently Reiss and Rodaro [6] for every ideal language** L presented a strongly connected automaton $\mathscr{A}$ such that $Syn(\mathscr{A}) = L$. Their proof is non-trivial and technical. In the present paper we give simple constructive proof of the fact that every finitely generated ideal language L, i.e. $L = \Sigma^* U \Sigma^*$ for some finite set U, serves as the language of synchronizing words of some strongly connected automaton. Our constructions reveal interesting connections with classical objects from combinatorics on words.

2 Algorithms and automata constructions

Let Σ be a finite alphabet with $|\Sigma| > 1$. Let L be a finitely generated ideal language over Σ, i.e. $L = \Sigma^* S \Sigma^*$, where S is a finite set of words. In this section we construct a strongly connected synchronizing automaton for which $L = \Sigma^* S \Sigma^*$.

First recall some standard definitions and fix notation. A word u is a *factor* (*prefix*, *suffix*) of a word w, if $w = xuy$ ($w = uy$, $w = xu$ respectively) for some

** over an alphabet with at least two letters

$x, y \in \Sigma^*$. By *Fact*(w) we denote the set of all factors of w. The i^{th} letter of the word w is denoted by $w[i]$. The factor $w[i]w[i+1]\cdots w[j]$ is denoted by $w[i..j]$. By Σ^n ($\Sigma^{\leq n}$, $\Sigma^{\geq n}$) we denote the set of all words over Σ of length n (at most n, at least n respectively).

Note, that if a word $s \in S$ is a factor of some other word $t \in S$, then the word t may be deleted from the set S without affecting the ideal language, generated by S. Thus, we may assume, that the set S is *anti-factorial*, i.e. no word in S is a factor of another word in S.

2.1 Ideal language generated by Σ^n

Theorem 1. *Let $\Sigma = \{a, b\}$. There is unique up to isomorphism strongly connected synchronizing automaton $\mathscr{B}$ such that* $\mathrm{Syn}(\mathscr{B}) = \Sigma^{\geq n}$.

Proof. Consider De Bruijn graph for the words of length n. Recall that the vertices of this graph are the words of length n, and there is a directed edge from the vertex u to the vertex v, if $u = xs$ and $v = sy$ for some $s \in \Sigma^{n-1}$, $x, y \in \Sigma$. By labeling each edge $e = (u, v)$ by the last letter of v we obtain De Bruijn automaton. Its state set is $Q = \Sigma^n$, and transition function is defined in the following way: $xs\,.\,y = sy$ for $s \in \Sigma^{n-1}$, $x, y \in \Sigma$. De Bruijn automaton is known to be strongly connected. Thus it remains to verify that $\mathrm{Syn}(\mathscr{B}) = L$. It is easy to see that for an arbitrary word u of length at most n we have $Q\,.\,u = \Sigma^{n-|u|}u$. Hence for any word w of length n we have $|Q\,.\,w| = 1$, and for any word u of length less than n we have $|Q\,.\,u| > 1$. So, $\mathrm{Syn}(\mathscr{B}) = L$.

Let $\mathscr{C} = \langle Q, \Sigma, \delta \rangle$ be a strongly connected synchronizing DFA such that $\mathrm{Syn}(\mathscr{C}) = L$. Let us prove that $|Q| \leq 2^n$. Strong connectivity implies $Q\,.\,a \cup Q\,.\,b = Q$. By induction it is easy to see that $Q = \bigcup_{|w|=k} Q\,.\,w$. In particular, we have $Q = \bigcup_{|w|=n} Q\,.\,w$. Thus, $|Q| = |\bigcup_{|w|=n} Q\,.\,w| \leq \sum_{|w|=n} |Q\,.\,w| = 2^n$. The last equality follows from the fact that every word of length n synchronizes $\mathscr{C}$, so each $Q\,.\,w$ is a singleton. For the converse inequality $2^n \leq |Q|$ consider the DFA $\mathscr{C}_a$, obtained from $\mathscr{C}$ by removing all transitions corresponding to the action of b in $\mathscr{C}$. The word a^n synchronizes $\mathscr{C}$, so $\mathscr{C}_a$ contains no cycles but unique loop. So the automaton $\mathscr{C}_a$ has a tree-like structure as it is shown on Fig.1. Denote by s the state of $\mathscr{C}$ such that $s\,.\,a = s$. The state s is called *root* of the tree, and the states $p_1, p_2, \ldots, p_k$ having no incoming transitions labeled by a are called *leaves* of the tree. The *height* $h(p_i)$ of a vertex p_i is the length of the path from p_i to the root s. The height of the tree $h(\mathscr{C}_a)$ is the maximal height of its leaves. We have $h(\mathscr{C}_a) = n$. Indeed, if $h(\mathscr{C}_a) = h < n$, then we would have $Q\,.\,a^h = \{s\}$, meaning that $a^h \in \mathrm{Syn}(\mathscr{C})$, which is impossible.

Consider the set of leaves $H = Q \setminus Q\,.\,a = \{p_1, p_2, ..., p_k\}$. Since the DFA $\mathscr{C}$ is strongly connected, for each state p_ℓ in H there exists a state q_ℓ such that $q_\ell\,.\,b = p_\ell$. Thus $H \subseteq Q\,.\,b$. We show that H is exactly $Q\,.\,b$, meaning that $Q\,.\,a \cap Q\,.\,b = \varnothing$. Take a leaf of height n. Without loss of generality suppose it is p_1. Let q_1 be such that $q_1\,.\,b = p_1$. The word ba^{n-1} is synchronizing, so $Q\,.\,ba^{n-1} = \{q\}$ for some $q \in Q$. We have $q_1\,.\,ba^{n-1} = q$, and $q\,.\,a = s$ (see Fig.1). Suppose there is $\overline{p} \in Q\,.\,a \cap Q\,.\,b$. Then there is a state $\overline{q}$ such that $\overline{q}\,.\,b = \overline{p}$.

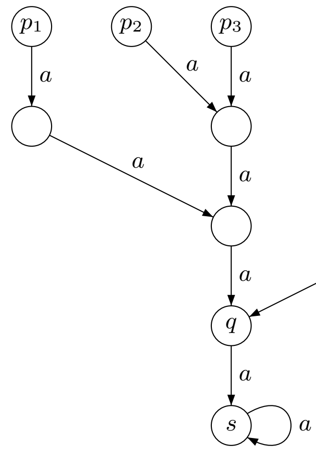

Fig. 1. The action of a in $\mathscr{C}$

Since $\overline{p}$ is not a leaf, we have $h(\overline{p}) < n$. Then $\overline{q} \,.\, ba^{n-1} = \overline{p} \,.\, a^{n-1} = s \neq q$. A contradiction. Hence $H = Q \,.\, b$. Furthermore, the height of any leaf of $\mathscr{C}_a$ is exactly n. To see this assume that there exists a state p_m such that $h(p_m) < n$, i.e. $p_m \,.\, a^\ell = s$, for some $\ell < n$. Then the word ba^{n-1} is not synchronizing. Indeed, take a state q_m such that $q_m \,.\, b = p_m$. We have $q_m \,.\, ba^{n-1} = p_m \,.\, a^{n-1} = s \neq q$.

Consider an arbitrary state $p \in Q \,.\, a$. Let $\delta^{-1}(p, u) = \{p' \in Q \mid p' \,.\, u = p\}$. We prove that $|\delta^{-1}(p, a)| \geq 2$ for each $p \in Q \,.\, a$. For the root s we have $\{s, q\} \subseteq \delta^{-1}(s, a)$, thus, $|\delta^{-1}(s, a)| \geq 2$. Let p be an arbitrary state in $Q \,.\, a$. Strong connectivity of $\mathscr{C}$ implies that there exists a state $\overline{p}$ and a word $w \in \Sigma^n$ such that $\overline{p} \,.\, w = p$. Since w is synchronizing, we have $Q \,.\, w = \{p\}$. Consider the word $w[1..n-1]$ that does not synchronize $\mathscr{C}$. Then $|Q \,.\, w[1..n-1]| \geq 2$. However, $(Q \,.\, w[1..n-1]) \,.\, w[n] = p$. And we obtain the inequality $|\delta^{-1}(p, a)| \geq 2$. Denote $H_0 = \{q\}$ and construct sets $H_i = \delta^{-1}(H_{i-1}, a)$ for $1 \leq i \leq n-1$. We have $|H_i| \geq 2^i$ for all $1 \leq i \leq n-1$. Then $\mathscr{C}$ possesses at least $1 + 1 + 2 + 4 + ... + 2^{n-1} = 2^n$ states.

Thus we have $|Q| = 2^n$. Moreover, $Q = \cup_{|w|=n} Q \,.\, w$. It means that with each state q of Q we can associate the word w of length n such that $Q \,.\, w = \{q\}$. It is clear that it gives us the desired isomorphism between $\mathscr{C}$ and $\mathscr{B}$.

□

Remark 1. In case $\Sigma = \{a, b\} \cup \Delta$, where $\Delta \neq \varnothing$, we consider De Bruijn automaton constructed for the binary alphabet $\{a, b\}$ and put the action of each letter in Δ to be the same as the action of the letter a. It is clear that the language of synchronizing words of the modified De Bruijn automaton coincides with $\Sigma^{\geq n}$.

The Proposition implies that the minimal DFA recognizing an ideal language L can be exponentially smaller than a strongly connected MSA $\mathscr{B}$ with $Syn(\mathscr{B}) = L$.

2.2 Ideal language generated by a set of words of fixed length

Theorem 2. *Let* $U \subsetneq \Sigma^n$. *There is a strongly connected synchronizing automaton* $\mathscr{B}_U$ *with* 2^n *states such that* $\mathrm{Syn}(\mathscr{B}_U) = \Sigma^* U \Sigma^*$.

Proof. We modify the De Bruijn automaton $\mathscr{B}$ from the section 2.1 to obtain the desired automaton $\mathscr{B}_U$. First of all it is convenient to view the states of the automaton $\mathscr{B}$ not as the words of length n, but as pairs (x, u), where $x \in \Sigma$ and $u \in \Sigma^{n-1}$. Then by the definition of the transitions in $\mathscr{B}$ we have

$$(x, u) \xrightarrow{y} (z, v) \Leftrightarrow uy = zv \tag{1}$$

For a word uy which is not in U, we modify the corresponding transition given by (1) in the following way. If $uy \notin U \cup \{a^n, b^n\}$ we put

$$(x, u) \xrightarrow{y} (x, v), \tag{2}$$

where v is defined by (1).

If $uy = a^n \notin U$ ($uy = b^n \notin U$ respectively) we put

$$(a, a^{n-1}) \xrightarrow{a} (b, a^{n-1}), \quad ((b, b^{n-1}) \xrightarrow{b} (a, b^{n-1}) \text{ respectively}). \tag{3}$$

The other transitions remain unchanged. The obtained automaton is denoted

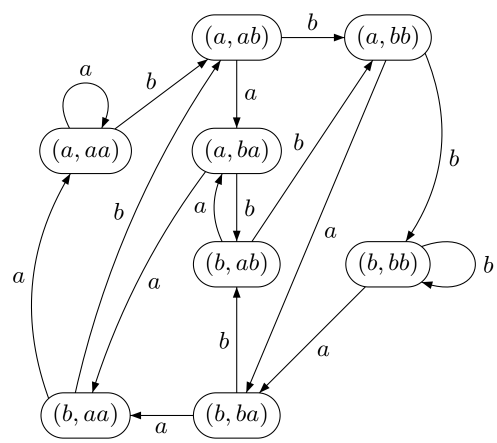

Fig. 2. De Bruijn automaton for $n = 3$

by $\mathscr{B}_U$. The examples of the automaton $\mathscr{B}$ and the corresponding modified automaton $\mathscr{B}_U$ for $U = \{aaa, abb, bab\}$ are shown on Fig.2 and Fig.3 respectively. We prove that the automaton $\mathscr{B}_U$ satisfies the statement of the proposition. First we show that $\mathscr{B}_U$ is strongly connected. For this purpose we prove that all the states are reachable from the state (a, a^{n-1}), and the state (a, a^{n-1}) is reachable from all states.

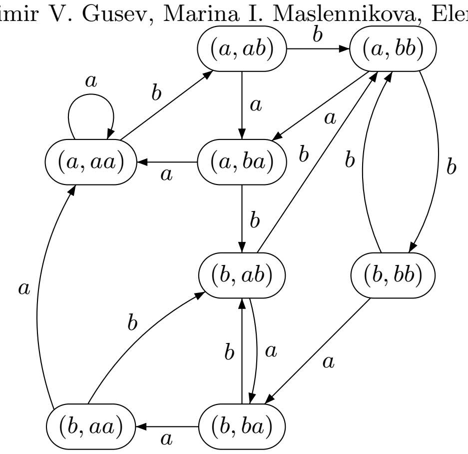

Fig. 3. Automaton $\mathscr{B}_U$ for $U = \{aaa, abb, bab\}$

First we show that a state (a, u) is reachable from (a, a^{n-1}) for any $u \in \Sigma^{n-1}$. If $u = a^{n-1}$, the claim obviously holds. Hence we may assume $u = a^k b\hat{u}$, where $k \geq 0$, $\hat{u} \in \Sigma^{n-k-2}$. By the definition of transitions in $\mathscr{B}_U$ we have

$$(a, a^{n-1}) \xrightarrow{b} (a, a^{n-2}b) \xrightarrow{\hat{u}[1]} (a, a^{n-3}b\hat{u}[1]) \xrightarrow{\hat{u}[2]} \cdots \xrightarrow{\hat{u}[n-k-3]}$$

$$(a, a^{k+1}b\hat{u}[1..n-k-3]) \xrightarrow{\hat{u}[n-k-2]} (a, a^k b\hat{u}[1..n-k-2]) = (a, u).$$

Symmetrically any state (b, u) is reachable from the state (b, b^{n-1}). The latter state is reachable from (a, b^{n-1}). Thus the state (b, u) is reachable also from (a, a^{n-1}):

$$(a, a^{n-1}) \rightsquigarrow (a, b^{n-1}) \xrightarrow{b} (b, b^{n-1}) \rightsquigarrow (b, u).$$

Now we show that the state (a, a^{n-1}) is reachable from any other state. Apply the word a^{n-1} to an arbitrary state (x, u). By the definition of transitions we have $(x, u) \,.\, a^{n-1} \in \{(a, a^{n-1}), (b, a^{n-1})\}$. If $(x, u) \,.\, a^{n-1} = (a, a^{n-1})$ we are done. If $(x, u) \,.\, a^{n-1} = (b, a^{n-1})$, then we apply once more the letter a and obtain $(x, u) \,.\, a^n = (a, a^{n-1})$.

Thus the constructed automaton $\mathscr{B}_U$ is strongly connected. Next we show that $\mathrm{Syn}(\mathscr{B}_U) = \Sigma^* U \Sigma^*$. It is easy to see that for any word $u \in \Sigma^{n-1}$ we have $Q \,.\, u \subseteq \{(a, u), (b, u)\}$, and $Q \,.\, u \cap Q \,.\, v = \varnothing$ for $u, v \in \Sigma^{n-1}$ such that $u \neq v$. Thus $Q \supseteq \bigcup\limits_{|u|=n-1} Q \,.\, u$. Next we check that $Q = \bigcup\limits_{|u|=n-1} Q \,.\, u$. Indeed, if $a^n \in U$ we have $(a, a^{n-1}) \xrightarrow{u} (a, u)$ for all $u \in \Sigma^{n-1}$. If $a^n \notin U$ take any word $u \in \Sigma^{n-1}$. If $u = a^{n-1}$ then u maps the state (a, a^{n-1}) or the state (b, a^{n-1}) to (a, u). Let us assume now that $u = a^k b\hat{u}$. If k is even (odd, respectively) then u maps (a, a^{n-1}) ((b, a^{n-1}), respectively) to (a, u). So any states (a, u) belongs to the set $\bigcup\limits_{|u|=n-1} Q \,.\, u$. Symmetrically any states (b, u) belongs to the latter set. Hence $Q = \bigcup\limits_{|u|=n-1} Q \,.\, u$. Since $|Q| = 2^n$, if there is a synchronizing word u of length

$n-1$, we would have $2^n = |Q| = |\bigcup_{|u|=n-1} Q\,.\,u| < 2^n$, which is a contradiction. Thus, none of the words of length $n-1$ is synchronizing. Consider an arbitrary word w of length n and factorize it as $w = uy$ with $u \in \Sigma^{n-1}$ and $y \in \Sigma$. We have $Q\,.\,u = \{(a,u),(b,u)\}$. If $w \in U$, then the corresponding transitions from the states (a,u) and (b,u) were not changed, and we have $Q\,.\,uy = \{(z,v)\}$, where $uy = zv$, so w is synchronizing. If $w \notin U$, then $Q\,.\,uy = \{(a,v),(b,v)\}$, where v is such that $uy = zv$ for some $z \in \Sigma$, so $w \notin \mathrm{Syn}(\mathscr{B}_U)$.

□

2.3 Ideal languages generated by a finite set of words

Theorem 3. *Let S be finite and anti-factorial set of words in Σ^+. There is a strongly connected synchronizing automaton $\mathscr{C}_S$ such that $\mathrm{Syn}(\mathscr{C}_S) = \Sigma^* S \Sigma^*$. This automaton has at most 2^n states, where $n = \max\{|s| \mid s \in S\}$.*

Proof. Let $T = \{w \in \Sigma^n \mid \exists s \in S, s \in \mathit{Fact}(w)\}$. First we construct the automaton $\mathscr{B}_T$ as described in the previous proposition. In that proposition the states of $\mathscr{B}_T$ were viewed as pairs (x,u) with $x \in \Sigma$, $u \in \Sigma^{n-1}$. Here it will be convenient to view the states as the words xu of length n (as it was in the initial De Bruijn automaton). Note, that since S is anti-factorial, every state in T can be uniquely factorized as usv such that $s \in S$, $u,v \in \Sigma^*$ and sv does not contain factors in S except s. In what follows we will use this unique representation without stating it explicitly.

Next we define an equivalence relation $\simeq$ on the set of states of this automaton (i.e. on words of length n) in the following way. Let $w, w' \in T$. We have $w \simeq w'$ iff $w = usv$ and $w' = u'sv$, where $s \in S$, $u, u', v \in \Sigma^*$. On the set $\Sigma^n \setminus T$ the relation $\simeq$ is defined trivially, i.e. for $w, w' \in \Sigma^n \setminus T$ we have $w \simeq w'$ iff $w = w'$. It is easy to see that $\simeq$ is indeed an equivalence relation on Σ^n. In fact, $\simeq$ is a congruence on the set of states of the automaton $\mathscr{B}_T$. Let us check that for any $x \in \Sigma$ and any $w, w' \in \Sigma^n$ $w \simeq w'$ implies $w\,.\,x \simeq w'\,.\,x$. If $w, w' \in \Sigma^n \setminus T$, then $w = w'$ and we are done. if $w, w' \in T$, then $w = usv$, $w' = u'sv$. If $u = u' = \varepsilon$, then $w = w'$, and there is nothing to prove. So we may assume, that $u, u' \neq \varepsilon$. Then $usv\,.\,x = tsvx$ and $u'sv\,.\,x = t'svx$ for some $t, t' \in \Sigma^*$. Since the obtained two words have the same suffixes, containing a word in S, they are equivalent. So we can consider the factor automaton $\mathscr{B}_T/\simeq$, whose states are the equivalence classes of $\simeq$, and the transition function is induced from the initial automaton. Let us denote by $[sv]$ the equivalence class of a word $usv \in T$, and by $[u]$ the equivalence class of a word $u \notin T$. We claim, that $\mathscr{C}_S = \mathscr{B}_T/\simeq$. In other words, the constructed automaton is strongly connected, and $\mathrm{Syn}(\mathscr{B}_T/\simeq) = \Sigma^* S \Sigma^*$. The first property holds trivially, since a factor automaton of a strongly connected automaton is strongly connected.

For any $w \in \Sigma^*$ and $s \in S$ previously in $\mathscr{B}_T$ we had $w\,.\,s = us$, where $u \in \Sigma^*$. Since S is anti-factorial, in $\mathscr{B}_T/\simeq$ we have $[w]\,.\,s = [s]$, so any $s \in S$ is synchronizing for the automaton $\mathscr{B}_T/\simeq$. Now let t be a synchronizing word, so there is a state $[w]$ such that for any state $[w']$ we have $[w']\,.\,t = [w]$. If $[w]$ is a one-element class, then the word t was synchronizing for the initial automaton $\mathscr{B}_T$, so t contains some word in S as a factor, i.e. $t \in \Sigma^* S \Sigma^*$. Consider the case where $[w]$ is a class consisting of elements $u_1sv,\ u_2sv, \ldots, u_ksv,\ k > 1$. Note that in this case $u_i \neq \varepsilon$ for each $i = 1, \ldots, k$. This means that $t = usv$ for some $u \in \Sigma^*$, thus, also in this case $t \in \Sigma^* S \Sigma^*$.

□

Complete this section with an example. Let $S = \{a^2, aba\}$ and $\Sigma = \{a, b\}$. Construct the corresponding set $T = \{a^3, a^2b, ba^2, aba\}$. Next build the DFA $\mathscr{B}_T$. The resulting automaton is shown on the left side of Fig.4.

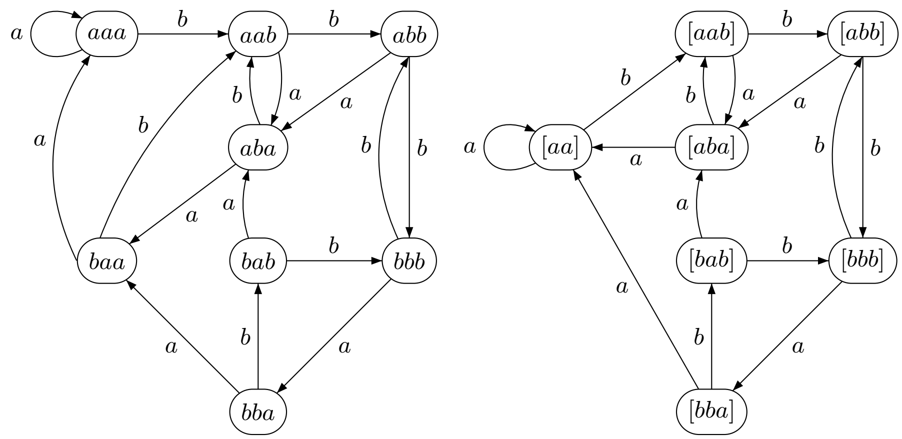

Fig. 4. Automata $\mathscr{B}_T$ and $\mathscr{B}_T/\simeq$ for $T = \{aaa, aab, baa, aba\}$

By the definition of $\simeq$ the class $[aa]$ includes states aaa and baa. The rest classes are one-element. The resulting automaton $\mathscr{B}_T/\simeq$ is shown on the right side of Fig.4

2.4 Ideal languages generated by two words

Let $S = \{u, v\} \subseteq \Sigma^+$ and let $|u| = n$, $|v| = m$. Again we suppose that S is anti-factorial. In this case we can construct a strongly connected automaton $\mathscr{D}_{u,v}$ such that $\mathrm{Syn}(\mathscr{D}_{u,v}) = \Sigma^*(u+v)\Sigma^*$ with $n+m$ states, thus improving construction from the previous section. For simplicity we state and prove the following theorem only for the case of binary alphabet, although the same argument works in general.

Theorem 4. *Let $\Sigma = \{a, b\}$, and let $u \in \Sigma^n \setminus \{ab^{n-1}, a^{n-1}b, ba^{n-1}, b^{n-1}a\}$, $v \in \Sigma^m \setminus \{ab^{m-1}, a^{m-1}b, ba^{m-1}, b^{m-1}a\}$. There is a strongly connected synchronizing automaton $\mathscr{D}_{u,v}$ having $n + m$ states such that* $\mathrm{Syn}(\mathscr{D}_{u,v}) = \Sigma^*(u + v)\Sigma^*$.

Proof. In order to obtain $\mathscr{D}_{u,v}$ we combine minimal automata for the languages $\Sigma^* u \Sigma^*$ and $\Sigma^* v \Sigma^*$. For a letter $x \in \{a, b\}$ by $\overline{x}$ we denote its *complementary* letter, i.e. $\overline{a} = b$, and $\overline{b} = a$. Recall the construction of the minimal automaton recognizing the language $\Sigma^* w \Sigma^*$, where $w \in \Sigma^+$. It is well-known that this automaton has $|w|+1$ states. We enumerate the states of this automaton by the prefixes of the word w so that the state $w[1..i]$ maps to the state $w[1..i+1]$ under the action of the letter $w[i+1]$ for all i, $0 \leq i < k$. The other letter $\overline{w[i+1]}$ sends the state $w[1..i]$ to state p such that p is the maximal prefix of w that appears in $w[1..i+1]$ as a suffix. The state w is the sink state of the automaton. The initial state is ε and the unique final state is w, see Fig.5 (the transitions labeled by complementary letters $\overline{w[i]}$ are not shown).

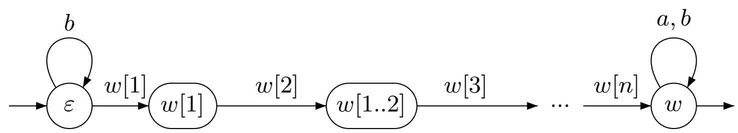

Fig. 5. The minimal DFA $\mathscr{A}_w$.

Construct minimal automata $\mathscr{A}_u$ and $\mathscr{A}_v$. Denote by $\mathscr{A}'_u$ the automaton obtained from $\mathscr{A}_u$ by deleting the sink state and the transition from $u[1..n-1]$ labeled by $u[n]$. Denote by $\mathscr{A}'_v$ the corresponding automaton for v. Define the action of letters $u[n]$ and $v[m]$ on states $u[1..n-1]$ and $v[1..m-1]$ as follows. Denote by p the state in $\mathscr{A}'_u$ corresponding to the maximal prefix of u that appears in v as a suffix. Denote by s the state in $\mathscr{A}'_v$ corresponding to the maximal prefix of v that appears in u as a suffix. We put $u[1..n-1] \,.\, u[n] = s$ and $v[1..m-1] \,.\, v[m] = p$. Denote the resulting automaton by $\mathscr{D}_{u,v}$ and prove that it satisfies the desired properties. Figures 6,7,8 illustrate the construction for $u = abaab$ and $v = babab$.

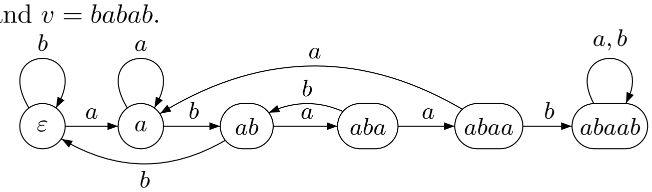

Fig. 6. The minimal DFA recognizing $\Sigma^* abaab \Sigma^*$.

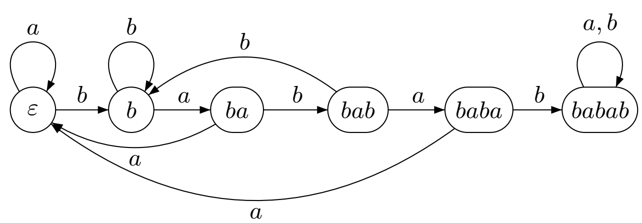

Fig. 7. The minimal DFA recognizing $\Sigma^* babab \Sigma^*$.

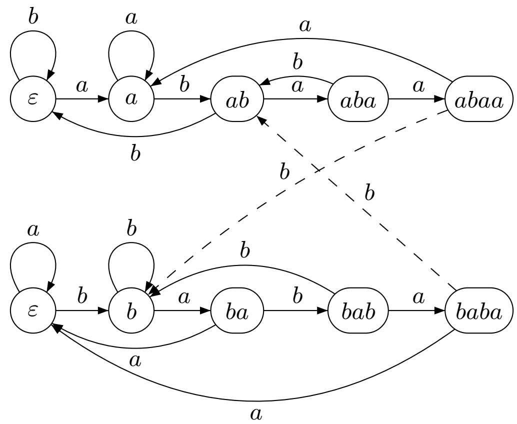

Fig. 8. The DFA $\mathscr{D}_{u,v}$.

The following claim is rather easy to see. The explicit proof can be found in [3].

Claim. If $w \in \Sigma^n \setminus \{a^{n-1}b, ab^{n-1}\}$, then the automaton $\mathscr{A}'_w$ is strongly connected.

By the Claim automata $\mathscr{A}'_u$ and $\mathscr{A}'_v$ are strongly connected. By the definition of the action of letters $u[n]$ and $v[m]$ on states $u[1..n-1]$ and $v[1..m-1]$, the resulting automaton $\mathscr{D}_{u,v}$ is also strongly connected.

Now we are going to verify that $u, v \in \mathrm{Syn}(\mathscr{D}_{u,v})$. The state set of $\mathscr{D}_{u,v}$ is the union of the state set of $\mathscr{A}'_u$ (denoted by Q^u), and the state set of $\mathscr{A}'_v$ (denoted by Q^v). To avoid confusion when necessary we will use the upper indices u and v for the states in Q^u and Q^v respectively. Let w be an arbitrary word. We claim that $\varepsilon^u . w = r$, where r is the maximal prefix of either u or v which is a suffix of w. Let us consider the path from ε^u to r. As long as we do not use modified transitions, i.e. the ones that lead from Q^u to Q^v or vice versa, the claim holds true by the definition of $\mathscr{A}'_u$ and $\mathscr{A}'_v$. Suppose now that the path contains a transition $u[1..n-1] \xrightarrow{u[n]} s$ and $\varepsilon^u . w' = u[1..n-1]$, where w' is a prefix of w. Let s' be the maximal prefix of the word v which is a suffix of $w'u[n]$. Note that $w'u[n]$ has u as a suffix. Therefore $|s'| < |u|$, otherwise u is a factor of v. Since $|s'| < |u|$ we have $s' = s$. Similar reasoning applies in case of transition $v[1..m-1] \xrightarrow{v[m]} p$. Therefore, the claim holds true. It is not hard to see that also $\varepsilon^v . w = r'$, where r' is the maximal prefix of either u or v which is a suffix of w.

Now we are ready to show that u is a synchronizing word for $\mathscr{D}_{u,v}$. By the definition of $\mathscr{D}_{u,v}$ we have $\varepsilon^u . u = s$. Let us consider an arbitrary state $t \in Q^u$. Let $\varepsilon^u . tu = r$. Note that the maximal prefix of the word u which is a suffix of tu is equal to u. Then by the claim r is a prefix of v. Since u is not a factor of v we have $|r| < |u|$. Thus, $r = s$ due to maximality. Since $\varepsilon^u . t = t$ we have $t . u = s$. Thus, $Q^u . u = \{s\}$. Arguing in the same way for the state ε^v we get $Q^v . u = \{s\}$. So, u is synchronizing. Analogously, one can show that v is synchronizing.

To complete the proof it remains to verify that each word from the set $\mathrm{Syn}(\mathscr{D}_{u,v})$ contains u or v as a factor. Take $w \in \mathrm{Syn}(\mathscr{D}_{u,v})$ and a state $r \in Q^u$. If $r \,.\, w \in Q^u$, then w maps all states in the component Q^v into the same state. In particular, $\varepsilon^v \,.\, w \in Q^u$. Thus v appears in w as a factor. Analogously if $r \,.\, w \in Q^v$, the word u appears as a factor in w. So we proved that $\mathrm{Syn}(\mathscr{D}_{u,v}) = \Sigma^*(u+v)\Sigma^*$. □

Acknowledgement. The authors acknowledge support from the Presidential Programm for young researchers, grant MK-266.2012.1.

References

1. J. Černý. *Poznámka k homogénnym eksperimentom s konečnými automatami.*, Mat.-Fyz. Cas. Slovensk. Akad. Vied. **14** (1964) 208–216.
2. Gusev V.V., Maslennikova M.I., Pribavkina E.V. *Principal ideal languages and synchronizing automata.* V. Halava, J. Karhumäki, Y. Matiyasevich, Proc. of the Second Russian Finnish Symposium on Discrete Mathematics, TUCS Lecture Notes, 2012, P. 79–84.
3. Gusev V.V., Maslennikova M.I., Pribavkina E.V. *Principal Ideal languages and synchronizing automata.*http://arxiv.org/abs/1304.3307(April 2013)
4. Maslennikova M.I. *Reset Complexity of Ideal Languages.* In M. Bieliková, G. Friedrich, G. Gottlob, S. Katzenbeisser, R. Špánek, G. Turán (eds.) Int. Conf. SOFSEM 2012, Proc. Volume II, Institute of Computer Science Academy of Sciences of the Czech Republic, 2012, P. 33–44.
5. D. Perrin *Finite automata.* Handbook of Theoretical computer Science, J. van Leewen, (ed.), Elsevier, B., P. 1–57, 1990.
6. R. Reis, E. Rodaro *On strongly connected ideal languages*, University of Porto Technical Report Series: DCC-2013-03.
7. S. Sandberg *Homing and synchronizing sequences*// In M. Broy et al (eds.) Model-Based Testing of Reactive Systems, Lect. Notes Comput. Sci, Springer-Verlag, Berlin-Heidelberg-New York. 2005. V.3472. P.5–33.
8. M. V. Volkov. *Synchronizing automata and the Černý conjecture*, in C. Martín-Vide, F. Otto, H. Fernau (eds.), Languages and Automata: Theory and Applications. LATA 2008. Lect. Notes Comp. Sci. **5196**, Berlin, Springer (2008) 11–27.