Regular Ideal Languages and Synchronizing Automata[⋆]

Rogério Reis and Emanuele Rodaro[⋆⋆]

Centro de Matemática, Universidade do Porto
R. Campo Alegre 687, 4169-007 Porto, Portugal
rvr@dcc.fc.up.pt, emanuele.rodaro@fc.up.pt

Abstract. We introduce the notion of reset left regular decomposition of an ideal regular language and we prove that there is a one-to-one correspondence between these decompositions and strongly connected synchronizing automata. We show that each ideal regular language has at least a reset left regular decomposition. As a consequence each ideal regular language is the set of synchronizing words of some strongly connected synchronizing automaton. Furthermore, this one-to-one correspondence allows us to formulate Černý's conjecture in a pure language theoretic framework.

1 Introduction

Since, in the context of this paper, we are not interested in automata as languages recognizer but just on the action of its transition function δ on the set of states Q, let us consider a deterministic finite automaton (DFA) as a tuple $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$, where the initial and final states are deliberately omitted from the definition. But, because in some point of this work we refer to an automaton as a language recognizer, we also call a DFA a tuple $\mathscr{B} = \langle Q', \Sigma', \delta', q_0, F \rangle$ and the language recognized by $\mathscr{B}$ is the set $L[\mathscr{B}] = \{u \in \Sigma^* : \delta'(q_0, u) \in F\}$. A DFA $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ is called synchronizing if there exists a word $w \in \Sigma^*$ "sending" all the states into a single one, i.e. $\delta(q, w) = \delta(q', w)$ for all $q, q' \in Q$. Any such word is said to be synchronizing (or reset) for the DFA $\mathscr{A}$. This notion has been widely studied since the work of Černý in 1964 [11] and his well known conjecture regarding the length of the shortest reset word. For more information on synchronizing automata we refer the reader to the survey by Volkov [12]. In what follows, when there is no ambiguity on the choice of the action δ of the automaton, we use the notation $q \cdot u$ instead of $\delta(q, u)$. We extend this action to a subset $H \subseteq Q$ in the obvious way $H \cdot u = \{q \cdot u : q \in H\}$ with the convention $\emptyset \cdot u = \emptyset$, and for a language $L \subseteq \Sigma^*$ we use the notation

[⋆] Work partialy supported by the European Regional Development Fund through the programme COMPETE and by the Portuguese Government through the FCT – Fundação para a Ciência e a Tecnologia under the project PEst-C/MAT/UI0144/2011 and CANTE-PTDC/EIA-CCO/101904/2008.

[⋆⋆] Partialy supported by FCT project SFRH/BPD/65428/2009.

J. Karhumäki, A. Lepistö, and L. Zamboni (Eds.): WORDS 2013, LNCS 8079, pp. 205–216, 2013.

$H \cdot L = \{q \cdot u : q \in H, u \in L\}$. We say that $\mathscr{A}$ is *strongly connected* whenever for any $q, q' \in Q$ there is a word $u \in \Sigma^*$ such that $q \cdot u = q'$. In the realm of synchronizing automata this notion is crucial since it is well known that Černý's conjecture is true if and only if it is true for the class of strongly connected synchronizing automata.

In this paper we study the relationship between ideal regular languages and synchronizing automata. A language $I \subseteq \Sigma^*$ is called a *two-sided ideal* (or simply an ideal) if $\Sigma^* I \Sigma^* \subseteq I$. In this work we will consider only ideal languages which are regular. Denote by $\mathbf{I}_\Sigma$ the class of ideal languages on an alphabet Σ. For a given synchronizing automaton $\mathscr{A}$, $\mathrm{Syn}(\mathscr{A})$ denotes the language of all the words synchronizing $\mathscr{A}$. It is a well known fact that $\mathrm{Syn}(\mathscr{A}) = \Sigma^* \,\mathrm{Syn}(\mathscr{A})\Sigma^*$ is a regular language which is also an ideal. This ideal is generated by the set of minimal synchronizing words $G = \mathrm{Syn}(\mathscr{A}) \setminus (\Sigma^+ \,\mathrm{Syn}(\mathscr{A}) \cup \mathrm{Syn}(\mathscr{A})\Sigma^+)$. This set can also be obtained considering the operators introduced in [6,8]. In case the set of generators G is finite, I is called finitely generated ideal and the synchronizing automata whose set of synchronizing words is finitely generated are called finitely generated synchronizing automata (see [5,7,9]). It is observed in [3] that the minimal deterministic automaton $\mathscr{A}_I = \langle Q', \Sigma, \delta', q_0, \{s\}\rangle$ recognizing an ideal language I is synchronizing with a unique final state s which is fixed by all the elements of Σ. We will refer to such state as *the sink state* for $\mathscr{A}_I$. Furthermore $\mathrm{Syn}(\mathscr{A}_I) = I$. Thus, each ideal language is endowed with at least a synchronizing automaton having I as the set of reset words. Therefore, for each ideal I there is a non-empty set $\mathcal{SA}(I)$ of all the synchronizing automata $\mathscr{B}$ with $\mathrm{Syn}(\mathscr{B}) = I$. In [3] the author introduces the notion of *reset complexity* of an ideal I as the number of states of the smallest automata in $\mathcal{SA}(I)$. In the same paper it is shown that the reset complexity can be exponentially smaller than the state complexity of the language. In [1] it is considered the special case of finitely generated synchronizing automata with the set of the reset words which is a principal ideal $P = \Sigma^* w \Sigma^*$ generated by a word $w \in \Sigma^*$, and it is presented an algorithm to generate a strongly connected synchronizing automaton $\mathscr{B}_w$ with $\mathrm{Syn}(\mathscr{B}_w) = P$ with the same number of states of $\mathscr{A}_P$. Therefore, for an ideal language I the first natural question that arises is wheather or not $\mathcal{SA}(I)$ always contains a strongly connected automaton or not. In Section 3 we answer affirmatively to this question for non-unary ideal languages. However, to study and characterize languages which are the reset words of strongly connected synchronizing automata we need to introduce the following provisional class of *strongly connected ideal language*:

Definition 1. *An ideal language I is called strongly connected whenever $I = \mathrm{Syn}(\mathscr{A})$ for some strongly connected synchronizing automaton $\mathscr{A}$.*

The paper is organized as follows. In Section 2 we introduce the notion of a (reset) left regular decomposition of an ideal, and we prove that strongly connected ideal languages are exactly the ideals having a reset left regular decomposition. We also exhibit a bijection that associates to each strongly connected ideal language I a strongly connected synchronizing automaton $\mathscr{A}$ with $\mathrm{Syn}(\mathscr{A}) = I$. In Section 3 we prove that each ideal language is a strongly connected ideal language. Thus, we can introduce the concept of reset regular decomposition complexity of an ideal and give an equivalent formulation of Černý's conjecture using this notion. Finally we state some open problems and direction of future research.

2 Strongly Connected Ideal Languages

We denote the class of strongly connected ideals on some finite alphabet Σ by $\mathbf{SCI}_\Sigma$ and the class of strongly connected synchronizing automata by $\mathbf{SCSA}_\Sigma$. Here, we characterize the class $\mathbf{SCI}_\Sigma$ using the concept of *reset left regular decomposition* of an ideal I. For $L \subseteq \Sigma^*$ and $u \in \Sigma^*$, let $Lu = \{xu : x \in L\}$, $uL = \{ux : x \in L\}$. The *reverse* operator $\cdot^R$ is such that given a word $u = u_1u_2 \ldots u_k$, $u^R = u_k \ldots u_2u_1$. This operator extends naturaly to languages.

Definition 2. *A left regular decomposition is a collection $\{I_i\}_{i\in F}$ of disjoint left ideals I_i of Σ^* for some finite set F such that:*

i) For any $a \in \Sigma$ and $i \in F$, there is a $j \in F$ such that $I_ia \subseteq I_j$.

The decomposition $\{I_i\}_{i\in F}$ is called a reset left regular decomposition if it also satisfies the following extra condition:

ii) Let $I = \uplus_{i\in F} I_i$. For any $u \in \Sigma^*$ if there is an $i \in F$ such that $Iu \subseteq I_i$, then $u \in I$.

Note that if $\{I_i\}_{i\in F}$ is a reset left regular decomposition, then the condition $Iu \subseteq I_i$ implies $u \in I_i$. Since $u \in I$, then $u \in I_j$ for some $j \in F$, hence $Iu \subseteq I_j$. If $j \neq i$ we have both $Iu \subseteq I_i$ and $Iu \subseteq I_j$ and thus $I_i \cap I_j \neq \emptyset$, which is a contradiction. We say that an ideal I has a (reset) left regular decomposition if there is a (reset) left regular decomposition $\{I_i\}_{i\in F}$ such that $I = \uplus_{i\in F} I_i$. The *order* of $\{I_i\}_{i\in F}$ is $|F|$. The notion of right regular decomposition is symmetric: exchange left ideals with right ideals and I_ia, Iu with aI_i, uI, respectively. Denote by $\mathbf{RLD}_\Sigma$ ($\mathbf{RRD}_\Sigma$) the class of the reset left (right) regular decompositions. Note that for a given left regular decomposition (reset left regular decomposition) $\{I_i\}_{i\in F}$, then $\{I_i^R\}_{i\in F}$ is a right regular decomposition (reset right regular decomposition). Thus $\cdot^R$ is a bijection between $\mathbf{RLD}_\Sigma \to \mathbf{RRD}_\Sigma$. We have the following characterization.

Theorem 3. *An ideal language I is strongly connected if and only if it has a reset left regular decomposition.*

Proof. Let $\mathscr{A} = \langle Q, \Sigma, \delta\rangle$ be a strongly connected synchronizing automata with $\mathrm{Syn}(\mathscr{A}) = I$. For each $q \in Q$, let:

$$I_q = \{u \in I : Q \cdot u = q\}$$

We claim that $\{I_q\}_{q\in Q}$ is a reset left regular decomposition for I. It is obvious that I_q are left ideals since for any $u \in I_q$ and $v \in \Sigma^*$, we get $Q\cdot vu \subseteq Q\cdot u = \{q\}$, i.e. $Q \cdot vu = \{q\}$. Let $q, q' \in Q$ with $q \neq q'$ and assume $I_q \cap I_{q'} \neq \emptyset$ and let

$u \in I_q \cap I_{q'}$. By definition, we have $q = Qu = q'$, which is a contradiction. Hence $I_q \cap I_{q'} = \emptyset$. Clearly $\uplus_{q\in Q} I_q \subseteq I$. Conversely if $u \in I$, since it is a reset word, then $Qu = q'$ for some $q' \in Q$, i.e. $u \in I_{q'}$ and so we have the decomposition $\uplus_{q\in Q} I_q = I$. Moreover for any $a \in \Sigma$, if $u \in I_q$, then $Q \cdot ua = q \cdot a$, thus $I_q a \subseteq I_{q\cdot a}$ and so condition i) of the Definition 2 is fulfilled. Thus it remains to prove that condition ii) is also satisfied. Suppose that $Iw \subseteq I_{\overline{q}}$ for some $\overline{q} \in Q$. Take any $q \in Q$, we claim that $qw = \overline{q}$ and so $w \in \mathrm{Syn}(\mathscr{A}) = I$. Take any $u' \in I$, thus $Q \cdot u' = q'$ for some $q' \in Q$. Since $\mathscr{A}$ is strongly connected, there is $u'' \in \Sigma^*$ such that $q' \cdot u'' = q$. Thus $u = u'u'' \in I$ satisfies $Q \cdot u = q$. Since $Iw \subseteq I_{\overline{q}}$ we get $\overline{q} = Q \cdot (uw) = q \cdot w$, i.e. $q \cdot w = \overline{q}$.

Conversely suppose that I has a reset left regular decomposition $\{I_i\}_{i\in F}$. We associate a DFA $\mathscr{A}(\{I_i\}_{i\in F}) = \langle \{I_i\}_{i\in F}, \Sigma, \eta\rangle$ in the following way. By condition i) of Definition 2 for any I_i and $a \in \Sigma$ there is a $j \in F$ with $I_i \cdot a \subseteq I_j$. Thus we define $\eta(I_i, a) = I_j$. This function is well defined. Let $j, k \in F$ with $j \neq i$, such that $I_i \cdot a \subseteq I_j, I_k$, then $I_i \cdot a \subseteq I_j \cap I_k$, hence $I_j \cap I_k \neq \emptyset$, which is a contradiction. Hence $\mathscr{A}(\{I_i\}_{i\in F})$ is a well defined DFA. It is straightforward to check that $\eta(I_i, u) = I_k$ for $u \in \Sigma^*$ if and only if $I_i u \subseteq I_k$. We prove that $\mathscr{A}(\{I_i\}_{i\in F})$ is strongly connected. Indeed take any $i, j \in F$ and let $w \in I_j$. Since I_j is a left ideal, then $I_i w \subseteq I_j$. Hence $I_i w \subseteq I_j$ implies $\eta(I_i, w) = I_j$ and so $\mathscr{A}(\{I_i\}_{i\in F})$ is strongly connected. We need to prove that $I \subseteq \mathrm{Syn}(\mathscr{A}(\{I_i\}_{i\in F}))$. Let $u \in I$, since $\{I_i\}_{i\in F}$ is a decomposition, $u \in I_j$ for some $j \in F$. Since I_j is a left ideal, we get $I_i u \subseteq I_j$ for any $i \in F$. Hence $\eta(I_i, u) = I_j$ for all $i \in F$, i.e. $u \in \mathrm{Syn}(\mathscr{A}(\{I_i\}_{i\in F}))$. Conversely, let $u \in \mathrm{Syn}(\mathscr{A}(\{I_i\}_{i\in F}))$. By the definition $\eta(I_i, u) = I_j$ for some $j \in F$ and for all $i \in F$. Therefore $I_i u \subseteq I_j$ which implies $Iu \subseteq I_j$ and so by ii) of Definition 2 we get $u \in I$. □

It is straightforward to check that the correspondence given in the proof of Theorem 3 is a bijection between the classes $\mathbf{RLD}_\Sigma$ and $\mathbf{SCSA}_\Sigma$. We state this fact in the following theorem.

Theorem 4. *The map $\mathcal{A} : \boldsymbol{RLD}_\Sigma \to \boldsymbol{SCSA}_\Sigma$ defined by*

$$\mathcal{A} : \{I_i\}_{i\in F} \mapsto \mathscr{A}(\{I_i\}_{i\in F}) = \langle \{I_i\}_{i\in F}, \Sigma, \eta\rangle$$

with $\eta(I_i, a) = I_j$ for $a \in \Sigma$ if and only if $I_i a \subseteq I_j$ is a bijection with inverse given by $\mathcal{I} : \boldsymbol{SCSA}_\Sigma \to \boldsymbol{RLD}_\Sigma$ defined by

$$\mathcal{I} : \mathscr{B} = \langle Q, \Sigma, \delta\rangle \mapsto \{I_q\}_{q\in Q} = \{\{u \in \Sigma^* : \delta(Q, u) = q\}\}_{q\in Q}$$

The following corollary characterizes the case of ideals on a unary alphabet.

Corollary 1. *Let I be an ideal over a unary alphabet $\Sigma = \{a\}$. Then I is strongly connected if and only if $I = \Sigma^*$.*

Proof. Since the alphabet is unary we have $I = a^* a^m a^*$ for some $m \geq 0$. Suppose that I is strongly connected, then by Theorem 3 there is a reset left regular decomposition $\{I_i\}_{i\in F}$ of I. Assume $a^m \in I_j$ for some $j \in F$. We claim $|F| = 1$. Indeed, since I_j is a left ideal we have $a^* a^m \subseteq I_j$, hence

$I = a^*a^ma^* = a^*a^m \subseteq I_j$, i.e. $I = I_j$. Therefore, by Theorem 4 the only strongly connected synchronizing automaton having I as set of reset words is the automaton with one state and a loop labelled by a. Hence $I = a^*$. On the other hand, if $I = a^*$ then I is the set of reset words of the synchronizing automaton with one state and a loop labelled by a, which is strongly connected, i.e. I is strongly connected. □

From this Corollary we can assume henceforth that the ideals considered are taken over an non-unary alphabet Σ. Given a strongly connected ideal language I with $\mathrm{Syn}(\mathscr{B}) = I$ for some strongly connected synchronizing automaton $\mathscr{B} = \langle Q, \Sigma, \delta\rangle$, there is an obvious way to calculate the associated reset left regular decomposition $\mathcal{I}(\mathscr{B})$. It is well known that I is recognized by the power automaton of $\mathscr{B}$ defined by $\mathcal{P}(\mathscr{B}) = \langle 2^Q, \Sigma, \delta, Q, \{\{q\} : q \in Q\}\rangle$, where 2^Q denotes the set of subsets of Q, the initial state is the set Q and the final set of states is formed by all the singletons. Thus, for each $q \in Q$ we can associate the DFA $\mathcal{P}(\mathscr{B})_q = \langle 2^Q, \Sigma, \delta, Q, \{q\}\rangle$ and so we can calculate the associated reset left regular decomposition by $\mathcal{I}(\mathscr{B}) = \{L[\mathcal{P}(\mathscr{B})_q]\}_{q\in Q}$. A first and quite natural issue is to calculate the reset left regular decompositions of the reset words of the Černý's series $\mathscr{C}_n = \langle\{1, \ldots, n\}, \{a, b\}, \delta_n\rangle$, where a acts like a ciclic permutation $\delta_n(i, a) = i + 1$ for $i = 1, \ldots, n-1$ and $\delta_n(n, a) = 1$, while b fixes all the states except the last one: $\delta_n(i, b) = i$ for $i = 1, \ldots, n-1$ and $\delta_n(n, b) = 1$ (see Fig. 1).

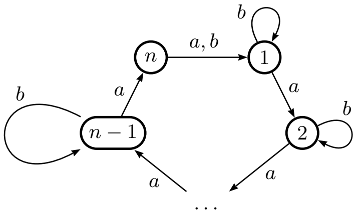

Fig. 1. The Černý's automaton $\mathscr{C}_n$

For example, in the case of $\mathscr{C}_4$ the associated reset left regular decomposition is the one given by

$$\begin{aligned} L[\mathcal{P}(\mathscr{C})_1] &= (((a^*b)(b + ab + a^4)^*(a^3b + (a^2b(b + a^2)^*ab)))((b + ab^*a^3) + \\ &\quad +((ab^*ab)(b + a^2)^*)ab))^*(ab^*a^2b)(b + ((ab^*ab^*)(a(a + b))))^* \\ L[\mathcal{P}(\mathscr{C})_2] &= L[\mathcal{P}(\mathscr{C})_1]ab^* \\ L[\mathcal{P}(\mathscr{C})_3] &= L[\mathcal{P}(\mathscr{C})_1]ab^*ab^* \\ L[\mathcal{P}(\mathscr{C})_4] &= L[\mathcal{P}(\mathscr{C})_1]ab^*ab^*a. \end{aligned}$$

In general, for $\mathscr{C}_n$ it is not difficult to see that $|\delta_n(\{1, \ldots, n\}, ux)| = 1$ and $|\delta_n(\{1, \ldots, n\}, u)| > 1$ for some word $u \in \{a, b\}^*$ and a letter $x \in \{a, b\}$ if and only if $\delta_n(\{1, \ldots, n\}, u) = \{n, 1\}$ and $x = b$. Thus, if $|\delta_n(Q, w)| = 1$, then there is a prefix $w'b$ of w with $\delta_n(Q, w') = \{n, 1\}$. Therefore, it is straightforward to check that in this case the decompositions are given by

$$\begin{aligned} L[\mathcal{P}(\mathscr{C})_1] &= \{w \in \Sigma^* \ : \ \delta_n(\{1, \ldots, n\}, w) = \{1\}\} \\ L[\mathcal{P}(\mathscr{C})_\ell] &= L[\mathcal{P}(\mathscr{C})_1](ab^*)^{\ell-1} \quad \text{for } \ell = 2, \ldots, n-1 \\ L[\mathcal{P}(\mathscr{C})_n] &= L[\mathcal{P}(\mathscr{C})_1](ab^*)^{n-2}a. \end{aligned}$$

By Theorem 3 if I is strongly connected, we can associate the non-empty set $\mathcal{R}(I)$ of all the reset left regular decompositions of I. We have the following lemma.

Lemma 1. *Let $\{I_i\}_{i\in F}$ be a reset left regular decompositions of I and let $\{J_k\}_{k\in H}$ be a left regular decomposition of an ideal J. If $I \subseteq J$, then the non-empty elements of $\{I_i \cap J_k\}_{i\in F, k\in H}$ form a reset left regular decomposition of I.*

Proof. Let $T \subseteq F \times H$ be the set of all the pairs of indices (i, j) for which $I_i \cap J_j \neq \emptyset$ and rename the set $\{I_i \cap J_k\}_{(i,k)\in T}$ by $\{S_j\}_{j\in T}$. It is clear that each S_j is a left ideal and $S_j \cap S_t = \emptyset$ for $j \neq t$. Furthermore $\uplus_{j\in T} S_j = I$. Condition i) is also verified. Take any S_j and suppose that $S_j = I_i \cap J_k$ for some $(i, k) \in T$, and let $a \in \Sigma$. Then $I_i a \subseteq I_s$, $J_k a \subseteq J_t$ for some $s \in F, t \in H$. Hence $(I_i \cap J_k)a = I_i a \cap J_k a \subseteq I_s \cap J_t = S_h$ for some $h \in T$, i.e. $S_j a \subseteq S_h$. Let us prove that reset condition ii) is also fulfilled. Assume $Iu \subseteq S_t$ for some $t \in T$ and $u \in \Sigma^*$. Thus $S_t = I_i \cap J_k$, for some $i \in F, k \in H$, hence $S_t \subseteq I_i$ which implies $Iu \subseteq I_i$. Hence $u \in I$ since $\{I_i\}_{i\in F}$ is a reset left regular decompositions of I. □

Given $\mathcal{I}, \mathcal{J} \in \mathcal{R}(I)$ with $\mathcal{I} = \{I_i\}_{i\in F}$ and $\mathcal{J} = \{J_k\}_{k\in H}$ by Lemma 1 the family $\mathcal{I} \wedge \mathcal{J} = \{I_i \cap J_k\}_{i\in F, k\in H}$ is still a reset left regular decomposition. Thus we have the following immediate result.

Corollary 2. *The family of the reset left regular decompositions of a strongly connected ideal I is a $\wedge$-semilattice.*

Let $\|I\| = \min\{|u| : u \in I\}$. It is a well known fact that Černý's conjecture holds if and only if it holds for strongly connected synchronizing automata. The following proposition place Černý's conjecture in a purely language theoretic context.

Proposition 5. *Černý's conjecture is true for strongly connected synchronizing automata if and only if for any strongly connected ideal I and any reset left regular decomposition $\{I_i\}_{i\in F}$ of I we have:*

$$|F| \geq \sqrt{\|I\|} + 1$$

Proof. Suppose that Černý's conjecture is true for strongly connected synchronizing automata. Let I be a strongly connected ideal and let $\{I_i\}_{i\in F}$ be a reset left regular decomposition of I. Let $\mathcal{A}(\{I_i\}_{i\in F})$ be the standard synchronizing automata associated to this decomposition as in Theorem 4. This automaton has $|F|$ states, hence there is a synchronizing word $u \in \mathrm{Syn}(\mathcal{A}(\{\mathscr{I}\}_{\in \mathscr{F}})) = I$ with $|u| \leq (|F| - 1)^2$. Thus $|F| \geq \sqrt{|u|} + 1 \geq \sqrt{\|I\|} + 1$.

Conversely, take any strongly connected synchronizing automata $\mathscr{A} = \langle Q, \Sigma, \delta \rangle$ with n states and let $\{I_q\}_{q \in Q}$ be the associated reset left regular decomposition of $I = \mathrm{Syn}(\mathscr{A})$ as in Theorem 4. Since the order of this decomposition is n, then $n \geq \sqrt{\|I\|} + 1$. Thus we have that there is a $u \in \mathrm{Syn}(\mathscr{A})$ with $|u| \leq (n-1)^2$ and so Černý's conjecture holds for $\mathscr{A}$. □

3 Ideal Languages Are Strongly Connected Ideal Languages

The notion of strongly connected ideal languages ($\mathbf{SCI}_\Sigma$) has been introduced in Section 2 to study the relationship between strongly connected synchronizing automata and ideal languages. In this section we show that $\mathbf{SCI}_\Sigma = \mathbf{I}_\Sigma$. This is done by showing that each ideal language I has at least a reset left regular decomposition. Equivalently, by Theorem 4, I is the set of the reset words of some strongly connected synchronizing automata with the same number of states as the order of this decomposition. However, the construction presented in Theorem 6 provides a reset left regular decomposition for I^R which is in general a double exponential with respect to the state complexity of I^R, and this bound does not seem to be tight. Before we prove the main result of this section we introduce some notions which are crucial for the sequel. Let $\mathscr{C} = \langle Q, \Sigma, \delta \rangle$ be an automaton with n states and a sink state s. Note that for such an automaton $|Q \cdot u| = 1$ if and only if $Q \cdot u = \{s\}$. Fix a word $u \in \Sigma^*$ and a subset $H \subseteq Q$. Assume $u = u_1 \dots u_r$ for $u_1, \dots, u_r \in \Sigma$ and $r = |u|$. For $0 \leq i < j \leq r$ we use the standard notation $u[i,j]$ to indicate the factor $u_i u_{i+1} \dots u_j$ if $i > 0$, otherwise $u[0,j] = u_1 \dots u_j$ with the convention that $u[0,0] = \epsilon$ and $u[i,i] = u_i$ if $i > 0$. We introduce a function which is fundamental in the sequel. Let $m = \frac{n^2+n}{2} + 1$ and let $\mathbb{Z}_m$ be the ring of the integers modulo m. For an integer $t \geq 1$, $[2^Q]_t$ denotes the set of subsets of Q of cardinality t. Let $\mathbb{T}_t = \mathbb{Z}_m([2^Q]_t \uplus \Sigma)$ be the free $\mathbb{Z}_m$-module on $[2^Q]_t \uplus \Sigma$. Let $H \in [2^Q]_t$, $a \in \Sigma$ and $p \in \mathbb{Z}_m([2^Q]_t \uplus \Sigma)$. We denote by $p(H)$, $p(a)$ the coefficients in $\mathbb{Z}_m$ of p with terms H, a, respectively. Note that p can be decomposed as the sum of the two following terms

$$p\langle Q \rangle = \sum_{H \subseteq Q} p(H)H, \quad p\langle \Sigma \rangle = \sum_{a \in \Sigma} p(a)a$$

Fix an element $u \in \Sigma^*$ with $u = u_1 \dots u_r$ and $H \subseteq Q$ with $|H| > 1$. Let j be the biggest index $1 \leq j \leq r$ such that $|H \cdot u[1,j]| > 1$ and if $j < n$, then $|H \cdot u[1, j+1]| = 1$. The set $S = H \cdot u[1,j]$ is called the *last set* of (H, u). Let i be the index $1 \leq i \leq r$ such that $u[i,j]$ is the maximal factor of u with $|S| = |H \cdot u[0,k]|$ for all $i \leq k \leq j$. The *tail* of (H, u) is the element of $\mathbb{Z}_m([2^Q]_t \uplus \Sigma)$ with $t = |S| \geq 2$ defined by

$$\mathcal{T}(H,u) = \begin{cases} \sum_{k=i}^{j-1} \left(H \cdot u[0,k] + u[k+1,k+1]\right), \text{if } u[0,j] = u \\ \sum_{k=i}^{j} \left(H \cdot u[0,k] + u[k+1,k+1]\right), \text{otherwise.} \end{cases}$$

Consider the set $\mathbb{T} = \uplus_{t=2}^{n} \mathbb{T}_t$. For an element $\mathcal{T} \in \mathbb{T}_t$, the integer $t \geq 2$ is called *the index* of $\mathcal{T}$ and it is denoted by $\mathrm{Ind}(\mathcal{T})$. We give to $\mathbb{T}$ a structure of semigroup by introducing an internal binary operation $\diamond$ defined in the following way. Let $\mathcal{T}_1 \in \mathbb{T}_i, \mathcal{T}_2 \in \mathbb{T}_j$, then

$$\mathcal{T}_1 \diamond \mathcal{T}_2 = \begin{cases} \mathcal{T}_{\min\{i,j\}} & \text{if } i \neq j \\ \mathcal{T}_1 + \mathcal{T}_2 & \text{otherwise} \end{cases}$$

Note that $(\mathbb{T}, \diamond)$ has a graded structure with respect to the semilattice $([2,n], \min)$, i.e. $\mathbb{T}_i \diamond \mathbb{T}_j \subseteq \mathbb{T}_{\min\{i,j\}}$. Let $u \in \Sigma^*$, the *tail map* is the function $\tau_u : 2^Q \to \mathbb{T}$ defined by

$$\tau_u(H) = \begin{cases} \mathcal{T}(H, u) & \text{if } |H| > 1 \\ 0_n & \text{otherwise} \end{cases}$$

where 0_n is the zero of $\mathbb{T}_n$. The following lemma is a direct consequence of the definitions.

Lemma 2. *With the above notation for any $u, v \in \Sigma^*$ we have:*

$$\tau_{vu}(T) = \tau_v(T) \diamond \tau_u(T \cdot v)$$

We denote by $\mathrm{Hom}(A, B)$ the set of the maps $f : A \to B$. We have the following lemma.

Lemma 3. *Consider the map $\mu : \Sigma^* \to \mathrm{Hom}(2^Q, \mathbb{T})$ defined by $\mu(u) = \tau_u$, then $\mathrm{Ker}(\mu)$ is a left congruence on Σ^*.*

We are now ready to prove the main theorem of this section.

Theorem 6. *Let $I \subseteq \Sigma^*$ be an ideal language, then I is a strongly connected ideal language.*

Proof. Put $J = I^R$. Let $\mathscr{A}_J = \langle Q, \Sigma, \delta, q_0, \{s\} \rangle$ be the minimal DFA recognizing J and let μ be the map of Lemma 3 defined with respect to $\mathscr{A}_J$. We claim that the equivalence classes of the relation $\sim = (J \times J) \cap \mathrm{Ker}(\mu)$ form a reset right regular decomposition of J. By the definition of the map μ, $\mathrm{Ker}(\mu)$ has finite index, thus $\sim$ has also finite index. Since $J = \mathrm{Syn}(\mathscr{A}_J)$, for any $H \subseteq Q$ and $u \in J$ we have $H \cdot u = \{s\}$. Hence it is straightforward to check that $\tau_u = \tau_{uv}$ for any $v \in \Sigma^*$. Therefore the $\sim$-classes are right ideals and form a finite partition $\{J_i\}_{i \in F}$ of J. Furthermore, by Lemma 3, $\mathrm{Ker}(\mu)$ is a left congruences of Σ^*, and so, since J is an ideal, it is also a congruence on J, hence for any J_i and $a \in \Sigma$, we get $aJ_i \subseteq J_j$ for some $j \in F$. Thus condition i) of Definition 2 is satisfied and so $\{J_i\}_{i \in F}$ is a right regular decomposition. We claim that also condition ii) is satisfied. Assume, contrary to our claim, that there are $i \in F$ and $v \in \Sigma^* \setminus J$ such that $vJ \subseteq J_i$. Write $H = Q \cdot v$. Since $\mathrm{Syn}(\mathscr{A}_J) = J$ we get $|H| > 1$. Thus let $t = \min\{|H \cdot r| : r \in \Sigma^* \text{ and } H \cdot r \neq \{s\}\}$ and let $S \in \{H \cdot r : r \in \Sigma^* \text{ and } |H \cdot r| = t\}$. Let $x \in \Sigma^*$ such that $H \cdot x = S$ and let $u = vx$. Note that $u \in \Sigma^* \setminus J$, $uJ \subseteq J_i$ and $Q \cdot u = S$ with $|S| = t$. Since $\mathrm{Syn}(\mathscr{A}_J) = J$ and $\mathscr{A}_J$ is a synchronizing automaton with zero, then there is a synchronizing word $w \in J$ with $|w| < \frac{n^2+n}{2} + 1$ where $n = |Q|$ (see [10]). Let T' be the last set of (S, w) and let w' be the maximal prefix of w such that $S \cdot w' = T'$. Thus, there is a letter $a \in \Sigma$ such that $w'a$ is a prefix of w and $|T'a| = 1$. We consider two mutually exclusive cases.

i) Suppose $|T' \cdot b| = 1$ for any $b \in \Sigma$. It is not difficult to check that $\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'a)$. Since $|\Sigma| > 1$ consider a letter $b \in \Sigma$ with $b \neq a$. Since $Q \cdot uw' = T'$ and $|T' \cdot b| = 1$, we also have $\mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b)$. Since $uJ \subseteq J_i$ we have $uw, uw'bw \in J_i$ (being $w'bw \in J$). Hence we get

$$\mathcal{T}(Q, uw'a) = \mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b)$$

In particular we get $\mathcal{T}(Q, uw'a)\langle\Sigma\rangle = \mathcal{T}(Q, uw'b)\langle\Sigma\rangle$, from which it follows $a = b$, a contradiction.

ii) Thus, we can assume that there is a letter $b \in \Sigma$, such that $|T' \cdot b| > 1$. Since $uw, uw'bw \in J_i$ (being $w, w'bw \in J$), we have $\mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw)$. Hence, by Lemma 2 we have

$$\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b) \diamond \mathcal{T}(T, w)$$

with $T = T' \cdot b$. Since $|T'| = t$ is minimal and $|T| > 1$ we have $|T| = |T'| = t$, hence $\mathrm{Ind}(\mathcal{T}(Q, uw'b)) = \mathrm{Ind}(\mathcal{T}(T, w)) = t$. Therefore, by the previous equality and the definition of $\diamond$ we get

$$\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'bw) = \mathcal{T}(Q, uw'b) + \mathcal{T}(T, w)$$

In particular we have

$$\mathcal{T}(Q, uw)\langle Q\rangle = \mathcal{T}(Q, uw'bw)\langle Q\rangle = \mathcal{T}(Q, uw'b)\langle Q\rangle + \mathcal{T}(T, w)\langle Q\rangle \qquad (1)$$

Furthermore, T' is the last set of $(Q, uw'a)$ and uw' is the maximal prefix of $uw'a$ such that $T' = Q \cdot uw'$, since $|T'| = |T|$ we have that T is the last set of $(Q, uw'b)$ and $uw'b$ is the maximal prefix of $uw'b$ with $T = Q \cdot uw'b$. Thus, by the definition of tail we have $\mathcal{T}(Q, uw'a)\langle Q\rangle = \mathcal{T}(Q, uw'b)\langle Q\rangle$. We have already observed that $\mathcal{T}(Q, uw) = \mathcal{T}(Q, uw'a)$, hence by (1)

$$\mathcal{T}(T, w)\langle Q\rangle = 0 \qquad (2)$$

Let $0 = i_1 < i_2 < \ldots < i_\ell \leq |w|$ be the maximal set of indices such that $T = T \cdot w[0, i_j]$ for all $1 \leq j \leq \ell$. Therefore, by the definition of tail and (2) we have in particular

$$0 = \mathcal{T}(T, w)(T) = \ell \mod \frac{n^2+n}{2} + 1$$

Since $\ell \geq 1$ we have that ℓ is a multiple of $\frac{n^2+n}{2} + 1$. However $\ell \leq |w| < \frac{n^2+n}{2} + 1$, which is a contradiction.

Therefore $v \in J$ and this concludes the proof of the fact that $\{J_i\}_{i\in F}$ is a reset right regular decomposition. Hence $\{J_i^R\}_{i\in F}$ is a reset left regular decomposition and so by Theorem 3 I is a strongly connected ideal language. □

Corollary 3. *Let I be an ideal language on Σ such that I^R has state complexity n. Then there is a strongly connected synchronizing automata $\mathscr{B}$ with N states and* $\mathrm{Syn}(\mathscr{B}) = I$ *such that:*

$$N \leq m^{k2^n} \left(\sum_{t=2}^{n} m^{\binom{n}{t}} \right)^{2^n}$$

where $k = |\Sigma|$ and $m = \left(\frac{n^2+n}{2} + 1 \right)$.

This corollary shows a double exponential upper bound for the number of states of the associated strongly connected automaton with respect to the state complexity of the reverse of the ideal language. It is unknown by the authors whether this bound is tight or not. In [1], for instance, it is shown an algorithm that given a principal ideal $I = \Sigma^* w \Sigma^*$ with $|w| = n$ in inputs, it returns a strongly connected synchronizing automaton with $n+1$ states. Therefore in this case the bound is linear with respect to the state complexity of I^R. Even more recently in this volume [2], it is proven that in case I is finitely generated there is always a strongly connected synchronizing automaton with a number of states upper bounded by $2^{\|I\|}$, and this bound is tight. Similarly to [3], where the author has introduced the notion of reset complexity of an ideal I (indicated by $\mathrm{rc}(I)$) as the number of states of the smallest synchronizing automaton $\mathscr{A}$ with $\mathrm{Syn}(\mathscr{A}) = I$, we can also give a similar notion in the realm of strongly connected synchronizing automata/reset left regular decomposition. By Theorem 6 for any ideal languages I, the set $\mathcal{R}(I)$ of all the reset left regular decompositions of I is non-empty. Thus we can define the *reset regular decomposition complexity of* I as the integer

$$\mathrm{rdc}(I) = \min\{|F| : \{I_i\}_{i \in F} \in \mathcal{R}(I)\}$$

By the correspondence introduced in Theorem 3, $\mathrm{rdc}(I)$ is also the number of states of the smallest strongly connected synchronizing automaton with the set of reset words equal to I. Furthermore $\mathrm{rc}(I) \leq \mathrm{rdc}(I)$ holds. The importance of the index $\mathrm{rdc}(I)$ can be also understood by the following theorem where we present a purely language theoretic restatement of Černý's conjecture.

Theorem 7. *Černý's conjecture holds if and only if for any ideal language I we have:*

$$\mathrm{rdc}(I) \geq \sqrt{\|I\|} + 1$$

where $\|I\| = \min\{|w| : w \in I\}$.

Proof. This a consequence of the fact that Černý's conjecture holds if and only if it holds for strongly connected automata and Proposition 5. □

Note that using the well known upper bound $(n^3 - n)/6$ (see [4]) for the shortest reset word of a synchronizing automaton, we have the bound $\mathrm{rdc}(I) \geq \sqrt[3]{6\|I\|}$. In general, a natural issue would be the study of bounds for $\mathrm{rdc}(I)$ depending on the state complexity of I or I^R. As we have already observed, Corollary 3 gives an upper bound to $\mathrm{rdc}(I)$ with respect to the state complexity of I^R which is not known to be tight.

Open Problems

We list some open problems originated by the previous results. Fix an ideal language I.

1. Give a tight upper bound of $\mathrm{rdc}(I)$ with respect to the state complexity of I^R or I.
2. In case I is finitely generated is true that $\mathrm{rdc}(I) \geq \|I\|+1$? The same problem in case I is a principal ideal language has been raised in [1]. This would give a better bound for the shortest synchronizing word for the class of finitely generated synchronizing automata with respect to the bound obtained in [9].
3. The proof of Theorem 6 uses the minimal DFA recognizing I^R. Is there a proof using another automaton associated to I?
4. Recall that $\mathcal{R}(I)$ is the set of all the reset left regular decompositions of I and the order of a decomposition $\mathcal{I} \in \mathcal{R}(I)$ is just the cardinality $|\mathcal{I}|$. We denote by $\mathcal{R}_k(I)$ the set of reset left regular decompositions of I of order $k \geq 1$.

   A quite natural question is whether $\sup\{k \geq 1 : \mathcal{R}_k(I) \neq \emptyset\} = \infty$ or not? In particular, what is the case if we consider I in the class of finitely generated ideals or in the even smaller class of principal ideals? This last case answers to the question whether or not, given a principal ideal I, there can there can be an arbitrarily large strongly connected DFA $\mathscr{A}$ with $\mathrm{Syn}(\mathscr{A}) = I$.
5. By Theorem 3, a naive way to calculate $\mathcal{R}_k(I)$ can be accomplished by building all the strongly connected synchronizing automata with k states and checking if their set of reset words coincides with I. Thus, it is natural to ask whether there is a more "efficient" way to perform this task without passing from the construction of all the automata with k states.

Acknowledgments. The authors thank E. Pribavkina for pointing out the unary case alphabet in Corollary 1.

References

1. Gusev, V., Maslennikova, M., Pribavkina, E.: Principal ideal languages and synchronizing automata. In: Halava, V., Karhumaki, J., Matiyasevich, Y. (eds.) RuFiDimII. TUCS Lecture Notes, vol. 17 (2012)
2. Gusev, V.V., Maslennikova, M.I., Pribavkina, E.V.: Finitely generated ideal languages and synchronizing automata. In: Karhumäki, J., Lepistö, A., Zamboni, L. (eds.) WORDS 2013. LNCS, vol. 8079, pp. 143–153. Springer, Heidelberg (2013)
3. Maslennikova, M.: Reset complexity of ideal languages. In: Bieliková, M., Friedrich, G., Gottlob, G., Katzenbeisser, S., Špánek, R., Turán, G. (eds.) Proc. Int. Conf. SOFSEM 2012, vol. II, pp. 33–44. Institute of Computer Science Academy of Sciences of the Czech Republic (2012)
4. Pin, J.E.: On two combinatorial problems arising from automata theory. Ann. Discrete Math. 17, 535–548 (1983)

5. Pribavkina, E.V., Rodaro, E.: Finitely generated synchronizing automata. In: Dediu, A.H., Ionescu, A.M., Martín-Vide, C. (eds.) LATA 2009. LNCS, vol. 5457, pp. 672–683. Springer, Heidelberg (2009)
6. Pribavkina, E.V., Rodaro, E.: State complexity of prefix, suffix, bifix and infix operators on regular languages. In: Gao, Y., Lu, H., Seki, S., Yu, S. (eds.) DLT 2010. LNCS, vol. 6224, pp. 376–386. Springer, Heidelberg (2010)
7. Pribavkina, E.V., Rodaro, E.: Recognizing synchronizing automata with finitely many minimal synchronizing words is PSPACE-complete. In: Löwe, B., Normann, D., Soskov, I., Soskova, A. (eds.) CiE 2011. LNCS, vol. 6735, pp. 230–238. Springer, Heidelberg (2011)
8. Pribavkina, E.V., Rodaro, E.: State complexity of code operators. International Journal of Foundations of Computer Science 22(07), 1669–1681 (2011)
9. Pribavkina, E.V., Rodaro, E.: Synchronizing automata with finitely many minimal synchronizing words. Information and Computation 209(3), 568–579 (2011), `http://www.sciencedirect.com/science/article/pii/S0890540110002063`
10. Rystsov, I.: Reset words for commutative and solvable automata. Theoretical Computer Science 172(1-2), 273–279 (1997), `http://www.sciencedirect.com/science/article/pii/S0304397596001363`
11. Černý, J.: Poznámka k homogénnym eksperimentom s konečnými automatami. Mat.-Fyz. Čas. Slovensk. Akad. Vied. 14, 208–216 (1964) (in slovak)
12. Volkov, M.V.: Synchronizing automata and the Černý conjecture. In: Martín-Vide, C., Otto, F., Fernau, H. (eds.) LATA 2008. LNCS, vol. 5196, pp. 11–27. Springer, Heidelberg (2008)