https://doi.org/10.1016/j.tcs.2015.08.042

To appear in *Theoretical Computer Science*

On incomplete and synchronizing finite sets *

Arturo Carpi
Dipartimento di Matematica e Informatica,
Università degli Studi di Perugia,
via Vanvitelli 1, 06123 Perugia, Italy.
e-mail: carpi@dmi.unipg.it

Flavio D'Alessandro
Dipartimento di Matematica,
La Sapienza Università di Roma
Piazzale Aldo Moro 2, 00185 Roma, Italy.
e-mail: dalessan@mat.uniroma1.it

Abstract

This paper situates itself in the theory of variable length codes and of finite automata where the concepts of completeness and synchronization play a central role. In this theoretical setting, we investigate the problem of finding upper bounds to the minimal length of synchronizing words and incompletable words of a finite language X in terms of the length of the words of X. This problem is related to two well-known conjectures formulated by Černý and Restivo, respectively. In particular, if Restivo's conjecture is true, our main result provides a quadratic bound for the minimal length of a synchronizing pair of any finite synchronizing complete code with respect to the maximal length of its words.

Keywords: Černý conjecture, synchronizing automaton, incompletable word, synchronizing set, complete set

*This work was partly supported by MIUR project PRIN 2010/2011 "Automi e Linguaggi Formali: Aspetti Matematici e Applicativi".

1 Introduction

The concepts of completeness and synchronization play a central role in Formal Language Theory since they appear in the study of several problems on variable length codes and on finite automata [5]. According to a well-known result of Schützenberger, the property of completeness provides an algebraic characterization of finite maximal codes, which are the objects used in Information Theory to construct optimal sequential codings.

Let X be a set of words on an alphabet A and let X^* be its Kleene closure. The set X is *complete* if any word on the alphabet A is a factor of some word belonging to X^*, otherwise it is *incomplete*. In the latter case, any word which is factor of no word of X^* is said to be *incompletable in* X.

In [21], Restivo conjectured that a finite incomplete set X has always an incompletable word whose length is quadratically bounded by the maximal length of the words of X. Results on this problem have been obtained in [6, 17, 18, 21]. The property of synchronization plays a natural role in Information Theory where it is relevant for the construction of decoders that are able to efficiently cope with decoding errors caused by noise during the data transmission. A set X is *synchronizing* if there are two words u, v of X^* such that whenever $ruvs \in X^*$, $r, s \in A^*$, one has also $ru, vs \in X^*$. The pair of words (u, v) is called a *synchronizing pair of* X.

In the study of synchronizing sets, the search for synchronizing words of minimal length in a prefix complete code is tightly related to that of synchronizing words of minimal length for synchronizing complete deterministic automata and the celebrated Černý Conjecture [15] (see also [2, 3, 4, 7, 8, 9, 10, 11, 12, 15, 19, 20, 23] for some results on the problem). In particular, in [3] (see also [4]), Béal and Perrin have proved that a complete synchronizing prefix code X on an alphabet of d letters with n code-words has a synchronizing word of length $O(n^2)$.

In this paper we are interested in finding upper bounds to the minimal lengths of incompletable and synchronizing words of a finite set X in terms of the size of X.

We recall that the size of X is the parameter $\ell(X)$ defined as the maximal length of the words of X.

Let $\mathcal{L}$ be a class of finite languages. For all $n, d > 0$, we denote by $R_{\mathcal{L}}(n, d)$ the least positive integer r satisfying the following condition: any incomplete set $X \in \mathcal{L}$ on a d-letter alphabet such that $\ell(X) \leq n$ has an incompletable word of length r. Similarly, we denote by $C_{\mathcal{L}}(n, d)$ the least positive integer c satisfying the following condition: any synchronizing set $X \in \mathcal{L}$ on a d-letter alphabet such that $\ell(X) \leq n$ has a synchronizing pair (u, v) such that $|uv| \leq c$.

In this context, the main result of this paper provides a bridge between the parameters $R_{\mathcal{L}}(n,d)$ and $C_{\mathcal{L}}(n,d)$. More precisely, denoting by $\mathcal{F}$ and by $\mathcal{M}$ the classes of finite languages and of complete finite codes respectively, we show that, for all $n, d > 0$,

$$C_{\mathcal{M}}(n,d) \leq 2R_{\mathcal{F}}(n,d+1) + 2n - 2.$$

In particular, if Restivo's conjecture is true, the latter bound gives

$$C_{\mathcal{M}}(n,d) = O(n^2),$$

thus providing a quadratic bound in the size of the set for the minimal length of a synchronizing pair of a finite synchronizing complete code.

In the second part of the paper, we study the dependence of the parameters $R_{\mathcal{L}}(n,d)$ and $C_{\mathcal{L}}(n,d)$ upon the number of letters d of the considered alphabet, by showing that both the parameters have a low rate of growth. More precisely, we show that, for the class $\mathcal{L}$ of finite languages (resp., codes, prefix codes), we have

$$R_{\mathcal{L}}(n,d) \leq \left\lceil \frac{R_{\mathcal{L}}(\lceil \log_2 d \rceil n, 2)}{\lfloor \log_2 d \rfloor} \right\rceil,$$

and, for the class $\mathcal{L}$ of finite complete languages (resp., codes, prefix codes), we have

$$C_{\mathcal{L}}(n,d) \leq \left\lceil \frac{C_{\mathcal{L}}(\lceil \log_2 (d+1) \rceil n, 2)}{\lfloor \log_2 (d-1) \rfloor} \right\rceil.$$

A similar result is obtained also when $\mathcal{L}$ is the class of finite (not necessarily complete) languages (resp., codes, prefix codes).

All the latter results were presented with a sketch of the proof in [13, 14].

The paper is structured as follows. In Section 2, some basic results about complete and synchronizing codes as well as synchronizing automata and Černý Conjecture are given. In Section 3 we describe our main result. In Section 4, a study of the dependence of the parameters $R_{\mathcal{L}}(n,d)$ and $C_{\mathcal{L}}(n,d)$ from the number d of letters of the alphabet is presented. Finally, in Section 5, some open questions about Restivo Conjecture are formulated.

2 Preliminaries

In this section we shortly recall some basic results of the theory of automata and of the theory of codes which will be useful in the sequel and we fix the corresponding notation used in the paper. The reader can refer to [5, 16] for more details.

2.1 Complete and synchronizing sets

Let A be a finite alphabet and let A^* be the free monoid of words over the alphabet A. The identity of A^* is called the *empty word* and is denoted by ϵ. The *length* of a word $w \in A^*$ is the integer $|w|$ inductively defined by $|\epsilon| = 0$, $|wa| = |w| + 1$, $w \in A^*$, $a \in A$. Given $w \in A^*$ and $a \in A$, we denote by $|w|_a$ the number of occurrences of the letter a in w. For any finite set of words W we denote by $\ell(W)$ the maximal length of the words of W. The number $\ell(W)$ will be called the *size* of W. Given words $u, w \in A^*$, u is said to be a *factor* of w if $w = \alpha u \beta$, for some $\alpha, \beta \in A^*$. The set of all factors of w is denoted by $\mathrm{Fact}(w)$. Given a set of words W, the set of the factors of all the words of W is denoted by $\mathrm{Fact}(W)$. Similarly, given a word w, a word u is said to be a *prefix* of w if $w = u\beta$, for some $\beta \in A^*$. A set X is said to be *prefix* if no word of X is a prefix of another word of X.

Definition 1 Let X be a subset of A^*. A pair of words (r, s) is an X-*completion* of a word w if $rws \in X^*$. A word having an X-completion is a *completable* word of X; conversely, a word with no X-completion is an *incompletable* word of X. The set X is *complete* if all words of A^* are completable words of X; X is *incomplete*, otherwise.

Another crucial notion of this paper is that of synchronizing set.

Definition 2 Let X be a subset of A^*. A pair $(u, v) \in X^* \times X^*$ is a *synchronizing pair* of X if for every X-completion (r, s) of uv, one has

$$ru, vs \in X^* .$$

The set X is *synchronizing* if it has a synchronizing pair.

Example 1 Consider the set

$$X = \{aa, ab, ba, baa, bbb\}$$

on the alphabet $A = \{a, b\}$. The pair (b, aa) is a X-completion of the word $bbabb$. Indeed, one has $b\,bbabb\,aa \in X^*$.

One easily verifies that all words of A^* of length 6 have an X-completion. On the contrary, the word $v = abbabba$ has no X-completion. Thus, v is an incompletable word of X of minimal length.

It is not difficult to verify that the pair (ab, ba) is a synchronizing pair of the set X. Thus, X is a synchronizing set.

The notion of synchronizing pair of a set is strictly related to that of *constant*. A word c of X^* is said to be a *constant* of X if, for every $u_1, u_2, u_3, u_4 \in A^*$ such that $u_1cu_2, u_3cu_4 \in X^*$, one has $u_1cu_4, u_3cu_2 \in X^*$. The following result holds.

Lemma 1 *Let X be a subset of A^*. If (u, v) is a synchronizing pair of X, then uv is a constant of X. Conversely, if c is a constant of X, then (c, c) is a synchronizing pair of X.*

2.2 Complete and synchronizing codes

The notions of complete and synchronizing sets provide a rich structure in the case that the set is a code. It is worth to shortly describe some fundamental results on such sets. A set X of words over an alphabet A is said to be a *(variable length) code over* A if it fulfills the unique factorization property, that is, for every word $u \in X^*$, there exists a unique sequence $x_1, \ldots, x_k$ of words of X such that $u = x_1 \cdots x_k$. A well-known example of codes is given by all prefix set which are distinct from $\{\epsilon\}$.

The notion of code is strictly related to the one of *monomorphism* of free monoids. Indeed, let A and B be two alphabets. As is well known, a morphism $h : A^* \to B^*$ is injective if and only if the letters of A have distinct images and $h(A)$ is a code.

In the sequel, a monomorphism $h : A^* \to B^*$ such that $h(A)$ is a prefix code will be called *prefix encoding.*

The notion of complete code is related to that of maximal code. Indeed, a regular code X is complete if and only if it is maximal (that is, it is not a subset of another code on the same alphabet). Moreover, a prefix code Y on an alphabet A is complete if and only if any word of A^* is a prefix of a word of X^* (see, *e.g.*, [5]).

2.3 Synchronizing automata and the Černý conjecture

As usually, by *finite non-deterministic automaton* we mean a 5-tuple $\mathcal{A} = \langle Q, A, \delta, I, F\rangle$, where Q is a finite set of elements called *states*, A is the *input alphabet*, $\delta : Q \times A \longrightarrow \mathcal{P}(Q)$ is the *transition function*, and $I, F \subseteq Q$ are the sets of initial and terminal states (here, $\mathcal{P}(Q)$ denotes the power set of Q).

With any automaton $\mathcal{A}$ is naturally associated a directed labelled finite multigraph $G(\mathcal{A}) = (Q, E)$, where the set E of edges is defined as

$$E = \{(p, a, q) \in Q \times A \times Q \mid q \in \delta(p, a)\}.$$

However, in this paper, we will consider only automata such that $I = F = \{1\}$, that is, with a unique initial and final state denoted 1. Such an automaton will be simply identified by the 4-tuple $\mathcal{A} = \langle Q, A, \delta, 1\rangle$. The language accepted by such an automaton is $L(\mathcal{A}) = X^*$, where X is the set of the labels of the paths in the graph $G(\mathcal{A})$, with origin and goal in the state 1, but with no intermediate vertex equal to 1.

The canonical extension of the map δ to the set $Q \times A^*$ will be still denoted by δ. Moreover, if P is a subset of Q and u is a word of A^*, we denote by $\delta(P, u)$ and $\delta(P, u^{-1})$ the sets:

$$\delta(P, u) = \{\delta(s, u) \mid s \in P\}, \quad \delta(P, u^{-1}) = \{s \in Q \mid \delta(s, u) \in P\}.$$

If no ambiguity arises, the sets $\delta(P, u)$ and $\delta(P, u^{-1})$ are denoted Pu and Pu^{-1}, respectively.

An automaton $\mathcal{A} = \langle Q, A, \delta, 1\rangle$ is said to be *transitive* if the graph $G(\mathcal{A})$ is strongly connected. It is not difficult to verify that any automaton $\mathcal{A}$ is equivalent to a transitive automaton whose graph is the strongly connected component of $G(\mathcal{A})$ containing the state 1. For this reason, in the sequel, we will consider only transitive automata.

An automaton $\mathcal{A} = \langle Q, A, \delta, 1\rangle$ is said to be *unambiguous* if for all $u, v \in A^*$ there is at most one state $q \in Q$ such that $q \in \delta(1, u)$ and $1 \in \delta(q, v)$. This is equivalent to say that any word of $L(\mathcal{A})$ is the label of a unique path of $G(\mathcal{A})$ with origin and goal in the state 1.

We say that an unambiguous automaton $\mathcal{A} = \langle Q, A, \delta, 1\rangle$ is *synchronizing* if there exist two words $w_1, w_2 \in A^*$ such that $Qw_1 \cap Qw_2^{-1} = \{1\}$.

The automaton $\mathcal{A}$ is *deterministic* if for all $q \in Q$ and for all $a \in A$, $\mathrm{Card}(qa) \leq 1$.

The automaton $\mathcal{A}$ is *complete* if for all $u \in A^*$, the set Qu is non-empty.

The properties of automata defined above reflects some properties of the minimal generating set X of the accepted language X^*. Some of them are summarized in the following lemma.

Lemma 2 *Let $X \subseteq A^*$ be the minimal generating set of X^* (that is, $X \cap X^2X^* = \emptyset$).*

1. *The set X is a regular code if and only if X^* is accepted by an unambiguous automaton $\mathcal{A} = \langle Q, A, \delta, 1\rangle$.*

2. *The set X is a prefix code if and only if X^* is accepted by a deterministic automaton $\mathcal{A} = \langle Q, A, \delta, 1\rangle$.*

3. *The set X is incomplete if and only if X^* is accepted by a transitive incomplete automaton $\mathcal{A} = \langle Q, A, \delta, 1\rangle$. Moreover, in such a case, a word $w \in A^*$ has an X-completion if and only if $Qw \neq \emptyset$.*

4. *The set X is a regular synchronizing code if and only if X^* is accepted by a transitive synchronizing unambiguous automaton $\mathcal{A} = \langle Q, A, \delta, 1\rangle$.*

Moreover, in such a case, a pair $(u, v) \in X^* \times X^*$ is a synchronizing pair of X if and only if $Qu \cap Qv^{-1} = \{1\}$.

As is well known, a deterministic automaton $\mathcal{A}$ is synchronizing if and only if there is a word u such that the set Qu is reduced to a single state.

Such a word is said to be a *synchronizing word* of $\mathcal{A}$. The following celebrated conjecture has been raised in [15].

Černý Conjecture. *Each synchronizing and complete deterministic automaton with n states has a synchronizing word of length $(n-1)^2$.*

Let us recall an important problem related to the Černý Conjecture. Let G be a finite, directed multigraph with all its vertices of the same outdegree. Then G is said to be *aperiodic* if the greatest common divisor of the lengths of all cycles of the graph is 1. The graph G is called a *AGW-graph* if it is aperiodic and strongly connected. The reason why such graphs take this name is due to the fact that these structures were first introduced and studied in the context of Symbolic Dynamics by Adler, Goodwyn and Weiss in [1].

A *synchronizing coloring* of G is a labeling of the edges of G that transforms it into a complete, deterministic and synchronizing automaton. The *Road coloring problem* asks for the existence of a synchronizing coloring for every AGW-graph. In 2007, Trahtman proved the following remarkable result [22].

Theorem 1 *Every AGW-graph has a synchronizing coloring.*

We recall that by the well known Kraft-McMillan Theorem (see, *e.g.*, [5]), integers $k_1, \ldots, k_n, d > 0$ are the code-word lengths of a maximal (or, equivalently, complete) prefix code over d letters if and only if they satisfy the condition

$$\sum_{i=1}^{n} d^{-k_i} = 1. \tag{1}$$

We conclude this section with an application of Trahtman Road-coloring Theorem, which furnishes a characterization of the code-word lengths of finite complete synchronizing codes.

Proposition 1 *Let $k_1, \ldots, k_n, d > 0$ be such that*

$$\gcd(k_1, k_2, \ldots, k_n) = 1\,, \qquad \sum_{i=1}^{n} d^{-k_i} = 1.$$

Then $k_1, \ldots, k_n$ are the code-word lengths of a synchronizing complete prefix code over d letters.

Proof Let A be a d-letter alphabet. By Kraft-McMillan Theorem, there exists a prefix code $X = \{x_1, \ldots, x_n\}$ over A such that, for every $i = 1, \ldots, n$, $|x_i| = k_i$. Moreover, such a code is maximal and, consequently, complete.

By Lemma 2, X^* is accepted by a complete deterministic automaton $\mathcal{A}_X$.

Let G be the underlying graph of $\mathcal{A}_X$, *i.e.*, the graph obtained from $\mathcal{A}_X$ by ripping off all the labels of its edges. Since $\gcd(k_1, k_2, \ldots, k_n) = 1$, G is an AGW-graph. By Theorem 1, there exists a synchronizing coloring $\mathcal{A}'$ of G. Let L be the language recognized by $\mathcal{A}'$. Again by Lemma 2, $L = Y^*$ for a suitable prefix complete synchronizing code Y. Moreover, by construction, one has $Y = \{y_1, \ldots, y_n\}$ with $|y_i| = |x_i| = k_i$ for every $i = 1, \ldots, n$, $|y_i| = k_i$.

□

Remark 1 It is worth noticing that the code-word lengths of any finite synchronizing complete code over d letters satisfies both the conditions of Proposition 1.

Indeed, as a straightforward consequence of Kraft-McMillan Theorem, the second condition is verified by any maximal (or, equivalently, complete) finite prefix code over d letters.

In order to verify the first one, let X be a finite synchronizing complete code, $(u, v) \in X^* \times X^*$ be a synchronizing pair of X, $a \in A$ be a letter, and (r, s) be an X-completion of the word $uvauv$. Then, one has $ruvauvs \in X^*$ and, consequently, $ru, vau, vs \in X^*$. One derives that the greatest common divisor m of the code-word lengths of X has to divide $|u|$, $|v|$, $|vau|$ and also $|vau| - |u| - |v| = 1$. Thus, $m = 1$.

3 The main result

The main result of this paper is related to a problem that was formulated in [21] by Restivo. Let $\mathcal{L}$ be a class of finite languages. For all $n > 0$ we set

$$R_{\mathcal{L}}(n) = \sup_{d \geq 1} R_{\mathcal{L}}(n, d), \quad C_{\mathcal{L}}(n) = \sup_{d \geq 1} C_{\mathcal{L}}(n, d).$$

In [21], it was conjectured that if $\mathcal{F}$ is the class of all finite languages, then $R_{\mathcal{F}}(n) \leq 2n^2$.If we restrict ourselves to prefix codes, we get

Proposition 2 ([21]) *Let $\mathcal{P}$ be the class of finite prefix codes. Then*

$$R_{\mathcal{P}}(n) \leq 2n^2.$$

However, in the general case, the previous bound was disproved in [17]. A more general and larger counterexample is given in [18]. We can thus state a slightly weaker version of the problem as follows.

Conjecture 1 (Restivo's Conjecture) Let $\mathcal{F}$ be the class of all finite languages. Then $R_{\mathcal{F}}(n) = O(n^2)$.

In this context, the main result of this paper is the following.

Proposition 3 *Let $\mathcal{M}$ be the class of complete finite codes. For all $n, d > 0$,*

$$C_{\mathcal{M}}(n, d) \leq 2R_{\mathcal{F}}(n, d + 1) + 2n - 2.$$

Before proving Proposition 3, it is convenient to discuss some interesting consequences of this result. First, if Restivo's conjecture is true, we get

$$C_{\mathcal{M}}(n) = O(n^2).$$

Moreover, the bound above would be sharp, as we explain below. Consider the prefix code $X_n = aA^{n-1} \cup bA^{n-2}$ on the alphabet $A = \{a, b\}$. The minimal automaton accepting X_n^* has been studied in [2], where it has been proved that the minimal length of its synchronizing words is $n^2 - 3n + 3$. From this, one derives that any synchronizing pair (w_1, w_2) of X_n verifies $|w_1 w_2| \geq (n-1)^2$. In particular, a synchronizing pair of X_n of minimal length is $((ab^{n-2})^{n-1}, \epsilon)$. This provides the lower bound

$$\mathcal{C}_{\mathcal{M}}(n, 2) \geq \mathcal{C}_{\mathcal{P}}(n, 2) \geq (n-1)^2,$$

for the parameter $\mathcal{C}_{\mathcal{M}}(n, 2)$.

It is also worth to do a remark on a recent result by Béal and Perrin. In [3] (cf. also [4]), it is proved that a synchronizing complete prefix code X with n code-words has a synchronizing word of length $2(n-2)(n-3)+1$. This result is derived from an upper bound to the length of shortest synchronizing words of synchronizing one-cluster automata. However, in view of Proposition 3 and Restivo's conjecture, this bound seems of no help in obtaining a good evaluation of the parameter $C_{\mathcal{P}}(n, 2)$, as one may have $n \simeq 2^{\ell(X)}$. This suggests that a bound in term of the size of X may be more informative than a bound in terms of the cardinality.

3.1 Proof of Proposition 3

Let us now proceed to prove Proposition 3. For this purpose, let X be a finite complete synchronizing code over a d-letter alphabet A and let $n = \ell(X)$. Let $\mathcal{A}_X = \langle Q, A, \delta, 1 \rangle$ be the unambiguous automaton that accepts X^* (see Lemma 2). The proof of Proposition 3 is based upon the following lemma.

Lemma 3 *Let (v_1, v_2) be a synchronizing pair of X. There exist words $w_1, w_2 \in A^*$ such that*

$$|w_1|, |w_2| \leq R_{\mathcal{F}}(n, d+1), \quad Qw_1 \subseteq Qv_1, \quad Qw_2{}^{-1} \subseteq Qv_2^{-1}.$$

Indeed, assume that Lemma 3 holds. As X is complete, the word w_1w_2 has an X-completion (r, s). With no loss of generality, we may suppose that $|r|, |s| \leq n-1$. Since (v_1, v_2) is a synchronizing pair, in view of Lemma 2, one has

$$Q(rw_1) \cap Q(w_2s)^{-1} \subseteq Qw_1 \cap Qw_2{}^{-1} \subseteq Qv_1 \cap Qv_2{}^{-1} = \{1\}.$$

Moreover, the word $rw_1w_2s \in X^*$ is accepted by $\mathcal{A}_X$ and therefore there is a state $q \in Q$ such that $q \in 1rw_1$ and $1 \in qw_2s$. Thus, $q \in Q(rw_1) \cap Q(w_2s)^{-1} \subseteq \{1\}$, that is, $q = 1$. This proves that $rw_1, w_2s \in X^*$ and by Lemma 2 (rw_1, w_2s) is a synchronizing pair of X. Moreover $|rw_1w_2s| \leq 2R_{\mathcal{F}}(n, d+1) + 2n - 2$. By the arbitrary choice of the maximal synchronizing code X, one derives Proposition 3.

Now, our main goal is to prove Lemma 3. For the sake of simplicity, we will prove the existence of the word w_1 that fulfills the conditions of Lemma 3 since the proof of the existence of the word w_2 can be obtained by using a symmetric construction. The main tool of this proof is a new automaton we construct below.

Let (v_1, v_2) be a synchronizing pair of X. If $v_1 = \epsilon$, the statement is trivially verified by $w_1 = v_1$. Thus we assume $v_1 \neq \epsilon$ and set $v_1 = ua$, with $u \in A^*$ and $a \in A$.

Let a' be a symbol not belonging to A and let $A' = A \cup \{a'\}$. We consider a new automaton $\mathcal{A}' = \langle Q, A', \delta', 1\rangle$ where the transition map δ' is defined as follows: for every $q \in Q$ and $a \in A$, $\delta'(q, a) = \delta(q, a)$ and

$$\delta'(q, a') = \begin{cases} \delta(q, a) \cup \{1\} & \text{if } q \notin \delta(Q, u), \\ \delta(q, a) \setminus \{1\} & \text{if } q \in \delta(Q, u). \end{cases} \tag{2}$$

It is useful to remark that, for all $q \in Q$ and for any word $w \in A^*$, $\delta'(q, w) = \delta(q, w)$. It is also useful to remark that, by construction, the automaton $\mathcal{A}'$ is still transitive. Let Y be the minimal generating set of the language accepted by $\mathcal{A}'$. Thus, $L_{\mathcal{A}'} = Y^*$ and $Y \cap Y^2Y^* = \emptyset$.

Now we prove some combinatorial properties of the set Y.

Lemma 4 *The set Y is incomplete.*

PROOF By (2) one has $\delta'(Q, ua') = \delta(Q, ua) \setminus \{1\} = \delta(Q, v_1) \setminus \{1\}$ and $\delta'(Q, v_2^{-1}) = \delta(Q, v_2^{-1})$. Taking into account that (v_1, v_2) is a synchronizing pair of X, one derives

$$\delta'(Q, ua') \cap \delta'(Q, v_2^{-1}) = \delta(Q, v_1) \cap \delta'(Q, v_2^{-1}) \setminus \{1\} = \emptyset\,.$$

It follows that $\delta'(Q, ua'v_2) = \emptyset$. This equation proves that the automaton $\mathcal{A}$ is not complete. Thus, by Lemma 2, Y is an incomplete set. □

Lemma 5 *It holds that $\ell(Y) \leq \ell(X)$.*

PROOF In order to prove the statement, it is enough to show that, for every $y \in Y$, there exists $x \in X$ with $|y| \leq |x|$.

Let $y = a_1 \cdots a_k \in Y$, with $a_i \in A'$, for $i = 1, \ldots, k$. Since $Y \cap Y^2Y^* = \emptyset$, in the graph of $\mathcal{A}'$ there is a path

$$c' = 1 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \xrightarrow{a_3} \cdots \xrightarrow{a_{k-1}} q_k \xrightarrow{a_k} 1,$$

where, for every $i = 1, \ldots, k$, $q_i \neq 1$. Let us now construct a path c in the graph of $\mathcal{A}_X$ such that $||c|| = x \in X$, with $|x| \geq |y|$, so completing the proof.

By the definition of $\mathcal{A}'$, any edge $p \xrightarrow{b} q$ of the graph of $\mathcal{A}'$ with $b \neq a'$ is also an edge of the graph of $\mathcal{A}$. Moreover, if $p \xrightarrow{a'} q$ is an edge of the graph of $\mathcal{A}'$ with $q \neq 1$, then $p \xrightarrow{a} q$ is an edge of the graph of $\mathcal{A}$. Thus, by replacing in c', every transition $q_i \xrightarrow{a'} q_{i+1}$, by $q_i \xrightarrow{a} q_{i+1}$ and deleting the last edge $q_k \xrightarrow{a_k} 1$, we find a path

$$d = 1 \xrightarrow{b_1} q_1 \xrightarrow{b_2} q_2 \xrightarrow{b_3} \cdots \xrightarrow{b_{k-1}} q_k \xrightarrow{b_k} 1,$$

of the graph of $\mathcal{A}$. Since $\mathcal{A}$ is transitive, one can catenate d with a simple path from q_k to 1. In such a way, we obtain a path c of the graph of $\mathcal{A}$ starting and ending in 1, with all intermediate states distinct from 1 and length $\geq k+1$. As is well known, as $\mathcal{A}$ is unambiguous, the label x of such a path is a word of the minimal generating set X of X^*. Since $|x| \geq k+1 = |y|$, this completes the proof. □

Lemma 6 *Let v be an incompletable word of Y of minimal length. There exists a word $w_1 \in A^*$ such that*

$$|w_1| \leq |v|, \quad Qw_1 \subseteq Qv_1.$$

PROOF Let v be an incompletable word of Y of minimal length, with the number $|v|_{a'}$ as small as possible. Then, by Lemma 2, one has $\delta'(Q,v)=\emptyset$.

The letter a' necessarily occurs in v, since by the completeness of $\mathcal{A}$, one has $\delta'(Q,r)=\delta(Q,r)\neq\emptyset$ for all $r\in A^*$. Thus, we can write $v=u_1a'u_2$, with $u_1\in A^*$ and $u_2\in A'^*$.

Recall that $v_1=ua$, with $u\in A^*$, $a\in A$. Let us verify that $\delta(Q,u_1)\subseteq\delta(Q,u)$. Indeed, suppose the contrary. Then, by (2), one has

$$\delta'(Q,u_1a')=\delta(Q,u_1a)\cup\{1\}=\delta'(Q,u_1a)\cup\{1\}$$

and consequently, $\delta'(Q,u_1au_2)\subseteq\delta'(Q,u_1a'u_2)=\emptyset$. Thus, u_1au_2 is an incompletable word of Y, but this contradicts the minimality of $|v|_{a'}$.

We conclude that $\delta(Q,u_1)\subseteq\delta(Q,u)$ and therefore taking $w_1=u_1a$ and recalling that $v_1=ua$, one has $\delta(Q,w_1)\subseteq\delta(Q,v_1)$ and $|w_1|\leq|v|$. The statement follows. □

Let us finally remark that Lemma 5 and Lemma 6 yield

$$|w_1|\leq R_{\mathcal{F}}(n,d+1),\quad Qw_1\subseteq Qv_1.$$

The proof of Lemma 3 is thus complete.

If we restrict ourselves to prefix codes, we obtain a tighter bound.

Proposition 4 *Let $\mathcal{MP}$ be the class of complete finite prefix codes. For all $n,d>0$,*

$$C_{\mathcal{MP}}(n,d)\leq R_{\mathcal{F}}(n,d+1)\,.$$

PROOF Let X be a maximal prefix code. Then, X is accepted by a complete deterministic automaton $\mathcal{A}_X$. Moreover, X has a synchronizing pair (v_1,v_2) with $v_2=\epsilon$. Thus, $Qv_1=Qv_1\cap Qv_2^{-1}=\{1\}$. By Lemma 3, there is a word $w_1\in A^*$ such that

$$|w_1|\leq R_{\mathcal{F}}(n,d+1)\,,\quad Qw_1=\{1\}\,.$$

This implies that $w_1\in X^*$ and (w_1,ϵ) is a synchronizing pair of the prefix code X. This proves the statement. □

Example 2 Consider the prefix code

$$X=\{a,baaa,baab,bab,bb\}\,.$$

The automata $\mathcal{A}_X$ and $\mathcal{A}'$ are represented in Figure 2. One obtains

$$Y=\{a,ba',bb,baa',bab,ba'a',ba'b,baaa,baab,\\ baa'a,baa'b,ba'aa,ba'ab,ba'a'a,ba'a'b\},$$

so that $\ell(Y)=\ell(X)=4$. The word aaa' is Y-incompletable and, consequently, (aaa,ϵ) is a synchronizing pair of the code X.

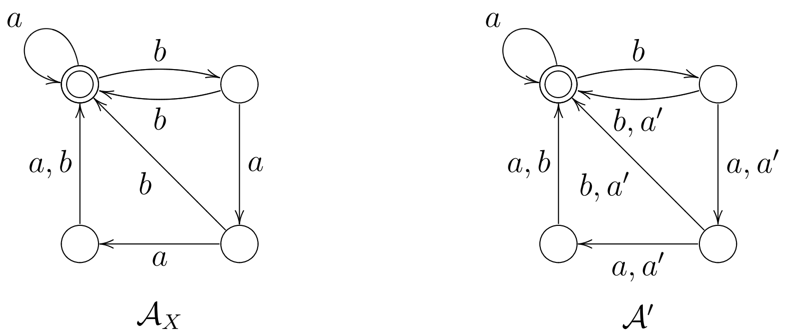

Figure 1: Automata of Example 2

4 Reduction to the binary case

The aim of this section is to study how much the parameters $R_{\mathcal{L}}(n, d)$ and $C_{\mathcal{L}}(n, d)$ vary according to the number d of letters of the alphabet. We start to analyze the parameter $R_{\mathcal{L}}(n, d)$. In the sequel, B denotes the binary alphabet $B = \{a, b\}$. The following lemma will be useful in the sequel. It gives an interesting insight on the structure of the completions of words in a complete regular set. As far as we know, it seems to be a new result.

Lemma 7 *Let $Y \subseteq A^*$ be a complete regular set. Then any word w of A^* has a Y-completion (y, s) with $y \in Y^*$.*

Proof We define an infinite sequence $((u_n, v_n))_{n \geq 0}$ as follows: (u_0, v_0) is a Y-completion of w; for all $n > 0$, (u_n, v_n) is a Y-completion of the word

$$wv_0wv_1 \cdots wv_{n-1}w.$$

By Myhill-Nerode Theorem (see, *e.g.*, [5]), Y^* is union of congruence classes of a congruence of finite index $\equiv$. Thus, one has $u_h \equiv u_k$ for some h, k with $k > h \geq 0$. By construction,

$$x = u_k wv_0wv_1 \cdots wv_k \in Y^* \quad \text{and} \quad z = u_h wv_0wv_1 \cdots wv_h \in Y^*.$$

One can write $x = yws$, with $y = u_k wv_0wv_1 \cdots wv_h$ and $s = v_{h+1}wv_{h+2} \cdots wv_k$, so that (y, s) is a Y-completion of w. Moreover, one has $y \equiv z$ and, consequently, $y \in Y^*$. This concludes the proof. □

Lemma 8 *Let $h : A^* \to B^*$ be a prefix encoding and $Y \subseteq A^*$. The set $h(Y)$ is complete if and only if Y and $h(A)$ are complete.*

Proof ($\Leftarrow$) Let $w \in B^*$. Since $h(A)$ is complete, one has $rws = h(u) \in h(A^*)$, for some $r, s \in B^*$ and $u \in A^*$. Since Y is complete, one has $puq \in Y^*$, where $p, q \in A^*$, thus yielding $h(puq) = h(p)rwsh(q) \in h(Y^*)$. Hence $(h(p)r, sh(q))$ is a $h(Y)$-completion of w.

($\Rightarrow$) The fact that $h(A)$ is complete follows straightforwardly from the inclusion $B^* \subseteq \mathrm{Fact}(h(Y^*)) \subseteq \mathrm{Fact}(h(A^*))$.

Let us prove that Y is complete. Let $w \in A^*$. Since $h(Y)$ is complete, by Lemma 7, one has $h(u)h(w)s = h(v)$, for some $u, v \in Y^*$ and $s \in B^*$. Since h is a prefix encoding, one has $v = uwr$, for some $r \in A^*$. The latter implies that (u, r) is a Y-completion of w. □

By encoding a d-letter alphabet on a suitable complete binary prefix code one obtains

Proposition 5 *Let $\mathcal{L}$ be the class of finite languages (resp., codes). Then*

$$R_{\mathcal{L}}(n, d) \leq \left\lceil \frac{R_{\mathcal{L}}(\lceil \log_2 d \rceil n, 2)}{\lfloor \log_2 d \rfloor} \right\rceil . \tag{3}$$

Proof Let A be a d-letter alphabet and let X be a finite incompletable language over A of size n. Set $m = \lceil \log_2 d \rceil$, $\gamma = 2^{m+1} - d$ and let $k_1, \ldots, k_d$ be the positive integers defined by

$$k_i = \begin{cases} m & \text{if } i \leq \gamma\,, \\ m+1 & \text{if } \gamma < i \leq d\,. \end{cases} \tag{4}$$

One easily checks that

$$\sum_{i=1}^{d} k_i = 1. \tag{5}$$

Thus, by Kraft-McMillan Theorem, $k_1, \ldots, k_d$ are the code-word lengths of a synchronizing prefix code Y over a binary alphabet B. Moreover, (5) ensures that Y is maximal and, consequently, complete.

Now, let $h : A^* \to B^*$ be a monomorphism such that $h(A) = Y$. Then, for every $a \in A$, we have

$$\lfloor \log_2 d \rfloor \leq |h(a)| \leq \lceil \log_2 d \rceil. \tag{6}$$

By (6) the size of $h(X)$ is not greater than $n\lceil \log_2 d \rceil$. By Lemma 8, since X is incompletable, $h(X)$ is incompletable as well. Let v be an incompletable word in $h(X)$ of minimal length. Hence we have

$$|v| \leq R_{\mathcal{L}}(\lceil \log_2 d \rceil n, 2). \tag{7}$$

Since $Y = h(A)$ is a complete prefix code, the word v is a prefix of a word of Y^*. Thus, $vs = h(u)$ for some $u \in A^*$ and $s \in B^*$. Moreover, taking u of minimal length, one may assume that $u = u'a$, with $u' \in A^*, a \in A$, and $|h(u')| < |v|$. In view of (6), one derives

$$|u| \leq \left\lceil \frac{|v|}{\lfloor \log_2 d \rfloor} \right\rceil . \tag{8}$$

Let us check that u is incompletable in X. By contradiction, deny. Then $r'us' \in X^*$, for some $r', s' \in A^*$. Consequently, $h(r'us') = h(r')vsh(s') \in h(X^*)$, thus implying that v is completable in $h(X)$.

Now (3) easily follows from the latter, (7) and (8). □

Let us now analyze the parameter $C_{\mathcal{L}}(n, d)$. The following lemma is useful for this purpose. It is algebraically similar to Lemma 8.

Lemma 9 *Let $h : A^* \to B^*$ be a monomorphism and let $Y \subseteq A^*$ be a complete set. The set $h(Y)$ is synchronizing if and only if Y and $h(A)$ are synchronizing.*

PROOF (⇐) By hypothesis and Lemma 1, there exists a word $y \in Y^*$ which is a constant of Y^*. Similarly, there exists a word $h(u) \in h(A^*)$, with $u \in A^*$, which is a constant of $h(A^*)$. Since Y is complete, there exist words $r, s \in A^*$ such that $rus \in Y^*$. Let $\zeta = h(rus) \in h(Y^*)$. Obviously, ζ is a constant of $h(A^*)$.

Let us prove that $\zeta h(y)\zeta \in h(Y^*)$ is a constant of $h(Y^*)$. For this purpose, let $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in B^*$ be such that $\alpha_1\zeta h(y)\zeta\alpha_2,\ \alpha_3\zeta h(y)\zeta\alpha_4 \in h(Y^*)$. Let us prove that (α_1, α_4) and (α_3, α_2) are $h(Y)$-completions of $\zeta h(y)\zeta$. By the latter condition and since ζ is a constant of $h(A^*)$, one has $\alpha_1\zeta \in h(A^*)$ so that $\alpha_1\zeta = h(\beta_1)$, for some $\beta_1 \in A^*$. Similarly, one has $\zeta\alpha_2 = h(\beta_2)$, $\alpha_3\zeta = h(\beta_3)$, $\zeta\alpha_4 = h(\beta_4)$, for some $\beta_2, \beta_3, \beta_4 \in A^*$. The previous two conditions now imply $h(\beta_1 y\beta_2), h(\beta_3 y\beta_4) \in h(Y^*)$. Since h is an injective map, the latter implies that $\beta_1 y\beta_2, \beta_3 y\beta_4 \in Y^*$. Since y is a constant of Y^*, one thus have $\beta_1 y\beta_4, \beta_3 y\beta_2 \in Y^*$ so that $h(\beta_1 y\beta_4), h(\beta_3 y\beta_2) \in h(Y^*)$, so implying that (α_1, α_4) and (α_3, α_2) are $h(Y)$-completions of $\zeta h(y)\zeta$.

(⇒) Let $(h(y_1), h(y_2))$ be a synchronizing pair of $h(Y)$, with $y_1, y_2 \in Y^*$. One easily proves that (y_1, y_2) is a synchronizing pair of Y. Indeed, if $ry_1y_2s \in Y^*$, with $r, s \in A^*$, one gets $h(ry_1y_2s) \in h(Y^*)$ which yields $h(ry_1), h(y_2s) \in h(Y^*)$. Since h is an injective map, from the latter we get $ry_1, y_2s \in Y^*$. Thus Y is a synchronizing set.

Let us prove now that $h(A)$ is a synchronizing set of B^* as well. More precisely, let us prove that the pair $(h(y_1), h(y_2))$ above considered, is a synchronizing pair of $h(A)$. For this purpose, let $r, s \in B^*$ such that $rh(y_1)h(y_2)s \in$

$h(A^*)$. Hence there exists $t \in A^*$ such that $rh(y_1y_2)s = h(t)$. On the other hand, since Y is complete, there exist words $t_1, t_2 \in A^*$ such that $t_1tt_2 \in Y^*$, which implies $h(t_1tt_2) = h(t_1)rh(y_1)h(y_2)sh(t_2) \in h(Y^*)$. Since $(h(y_1), h(y_2))$ is a synchronizing pair of $h(Y^*)$, one derives $h(t_1)rh(y_1)$, $h(y_2)sh(t_2) \in h(Y^*)$. Thus, one has

$$h(t_1),\ h(t_1)rh(y_1),\ rh(y_1)h(y_2)s,\ h(y_2)sh(t_2),\ h(t_2) \in h(A^*). \tag{9}$$

Taking into account that $h(A)$ is a code and, consequently, there is a unique factorization of the word $h(t_1)rh(y_1)h(y_2)sh(t_2)$ as product of words of $h(A)$, one derives

$$rh(y_1),\ h(y_2)s \in h(A^*)\,.$$

Hence, $(h(y_1), h(y_2))$ is a synchronizing pair of the code $h(A)$. This completes the proof. □

As an application of the two lemmas above, by encoding a d-letter alphabet on a suitable complete binary synchronizing code, one obtains the following result:

Proposition 6 *Let $\mathcal{L}$ be the class of finite complete languages (resp., codes, prefix codes). Then*

$$C_{\mathcal{L}}(n, d) \le \left\lceil \frac{C_{\mathcal{L}}(\lceil \log_2(d+1) \rceil n, 2)}{\lfloor \log_2(d-1) \rfloor} \right\rceil. \tag{10}$$

Proof Let A be a d-letter alphabet and let X be a finite complete synchronizing language over A of size n.

First, we consider the case that d is not a power of 2. Set $m = \lfloor \log_2 d \rfloor$, $\gamma = 2^{m+1} - d$ and let $k_1, \ldots, k_d$ be the positive integers defined by (4). One easily checks that both the conditions of Proposition 1 are satisfied. Thus, $k_1, \ldots, k_d$ are the code-word lengths of a synchronizing complete prefix code Y over a binary alphabet B.

Now, let $h : A^* \to B^*$ be a monomorphism such that $h(A) = Y$. Then (6) is verified by every $a \in A$, so that the size of $h(X)$ is not greater than $n\lceil \log_2 d \rceil$. Since X is a synchronizing and complete set and Y is a synchronizing and complete code, by Lemma 8 and Lemma 9, one has that $h(X)$ is a synchronizing and complete set as well. Moreover, if X is a code (resp., a prefix code), then $h(X)$ is a code (resp., a prefix code), too.

Let $(h(u), h(v))$ be a synchronizing pair of $h(X)$, $u, v \in X^*$. Hence we have

$$|h(uv)| \le C_{\mathcal{L}}(n\lceil \log_2 d \rceil, 2). \tag{11}$$

It is easily checked that (u, v) is a synchronizing pair of X. Indeed, let $ruvs \in X^*$, with $r, s \in A^*$. Hence $h(ruvs) \in h(X^*)$ so that $h(ru), h(vs) \in h(X^*)$. Since h is an injective mapping, we conclude that $ru, vs \in X^*$.

Hence, by taking account of (6), (11), one gets (10).

Finally, let us treat the case where $d = 2^m$. Let $k_1, \ldots, k_d$ be the sequence of positive integers defined as: for every $i = 1, \ldots, d$,

$$k_i = \begin{cases} m-1 & \text{if } i = 1\,, \\ m+1 & \text{if } i = 2, 3\,, \\ m & \text{if } i = 4, \ldots, d\,. \end{cases}$$

As before, one easily checks that the sequence of lengths $k_1, \ldots, k_d$ defined above satisfy both the conditions of Proposition 1. Thus, $k_1, \ldots, k_d$ are the code-word lengths of a synchronizing complete prefix code Y over a binary alphabet B. Moreover, for every $a \in A$, we have

$$\lfloor \log_2(d-1) \rfloor \leq |h(a)| \leq \lceil \log_2(d+1) \rceil.$$

From that point on, one proceeds by using the same argument of the previous case. The proof of the statement is now complete. □

A similar bound can be found also in the case where completeness is not required:

Proposition 7 *Let $\mathcal{L}$ be the class of finite languages (resp. codes, prefix codes). Then*

$$C_{\mathcal{L}}(n, d) \leq \left\lceil \frac{C_{\mathcal{L}}(\lceil \log_2(d+1) \rceil n, 2)}{\lceil \log_2(d+1) \rceil} \right\rceil. \tag{12}$$

Proof Let $X \subseteq B^m$, with $m \geq 1$ such that $a^m \notin X$ and $a^{m-1}b, ba^{m-1} \in X$. It is easily checked that X is a prefix synchronizing code endowed with the synchronizing pair $(ba^{m-1}, a^{m-1}b)$. Let A be a d-letter alphabet and let Y be a synchronizing set over A such that $\ell(Y) \leq n$. We will find a synchronizing pair of Y.

We may suppose that $Y \not\subseteq a^*$ since otherwise it has a synchronizing pair (u, v) with $|uv| \leq C_{\mathcal{L}}(n, 1) \leq C_{\mathcal{L}}(n, 2)$. Let (y_1, y_2) be a synchronizing pair of Y. With no loss of generality, we may assume that $ab \in \mathrm{Fact}(y_1 y_2)$, for two suitable distinct letters a, b. Let $m = \lceil \log_2(d+1) \rceil$ and let us consider the monomorphism $h : A^* \to B^*$ generated by a bijective mapping between A and a subset of the set X defined above such that

$$h(a) = ba^{m-1}, \quad h(b) = a^{m-1}b.$$

Let us prove that $(h(y_1), h(y_2))$ is a synchronizing pair of $h(Y)$ so that $h(Y)$ is a synchronizing set. For this purpose, let $rh(y_1)h(y_2)s \in h(Y^*)$ with $r, s \in B^*$. By costruction, we know that $y_1y_2 = \alpha ab\beta$, where $\alpha, \beta \in A^*$. The latter implies that

$$rh(y_1)h(y_2)s = rh(\alpha)ba^{m-1}a^{m-1}bh(\beta)s \in X^* .$$

Since $(ba^{m-1}, a^{m-1}b)$ is a synchronizing pair of X and X is a uniform length code, from the latter equation one has $r, s \in X^*$ and thus $r = h(r')$ and $s = h(s')$ with $r', s' \in A^*$. Hence $rh(y_1)h(y_2)s = h(r'y_1y_2s') \in h(Y^*)$. By the injectivity of h, one has $r'y_1y_2s' \in Y^*$. Since (y_1, y_2) is a synchronizing pair of Y, one derives $r'y_1, y_2s' \in Y^*$ and thus $rh(y_1), h(y_2)s \in h(Y^*)$.

Now, using an argument similar to that used in the proof of Proposition 6 and by remarking that, for every $w \in A^*$, $|h(w)| = |w|\lceil\log_2(d+1)\rceil$, one proves (12). □

5 Conclusions

In this paper we have studied the minimal lengths of incompletable and synchronizing words of a finite set X in terms of the size of X. In particular, we have shown some relations among the parameters $R_{\mathcal{F}}(n, d)$ and $C_{\mathcal{M}}(n, d)$ bounding, respectively, the minimal lengths of incompletable words in sets of size n and the minimal lengths of synchronizing pairs in maximal codes of size n.

As we have seen, Restivo conjectured a quadratic bound to the minimal length of incompletable words of any finite incompletable set. However, up to now, such a bound has been found only for prefix codes. Thus, we may consider the following unanswered questions, most of which may be viewed as weaker versions of Restivo's Conjecture. We recall that with $\mathcal{F}$ we have denoted the class of all finite sets.

1. Does $R_{\mathcal{F}}(n) < \infty$ for all n holds true?
2. Find a polynomial upper bound to $R_{\mathcal{F}}(n)$.
3. Find a polynomial upper bound to $R_{\mathcal{F}}(n, 2)$.
4. Let $\mathcal{F}_k$ be the class of all k-word languages ($k \geq 2$). Evaluate $R_{\mathcal{F}_k}(n)$.
5. Does $R_{\mathcal{F}_k}(n) = R_{\mathcal{F}_k}(n, 2)$ holds true?
6. Let $\mathcal{C}$ be the class of finite codes. Find a polynomial upper bound to $R_{\mathcal{C}}(n)$.

7. Let $\mathcal{P}$ be the class of finite prefix codes. Find the exact value of $R_{\mathcal{P}}(n)$ for all n.

References

[1] R. L. Adler, L. W. Goodwyn, B. Weiss, Equivalence of topological Markov shifts, *Israel J. Math.,* 27, 49–63, 1977.

[2] D. S. Ananichev, V. V. Gusev, M. V. Volkov, Slowly synchronizing automata and digraphs, in: P. Hliněný, A. Kučera eds., *MFCS 2010 Mathematical Foundations of Computer Science, Lecture Notes in Comput. Sci.* Vol. 6281, pp. 55–65, Springer, Berlin, 2010.

[3] M.-P. Béal, D. Perrin, A quadratic upper bound on the size of a synchronizing word in one-cluster automata, in: V. Diekert, D. Nowotka eds., *DLT 2009 Developments in Language Theory, Lecture Notes in Computer Science*, Vol. 5583, pp. 81–90, Springer, Berlin, 2009.

[4] M.-P. Béal, M. V. Berlinkov, D. Perrin, A quadratic upper bound on the size of a synchronizing word in one-cluster automata, *Int. J. Found. Comput. Sci.,* 22, 277–288, 2011.

[5] J. Berstel, D. Perrin, Ch. Reutenauer, Codes and Automata, Encyclopedia of Mathematics and its Applications, 129, Cambridge University Press, 2009.

[6] J. M. Boë, A. de Luca, A. Restivo, Minimal complete sets of words, *Theoret. Comput. Sci.,* 12, 325–332, 1980.

[7] A. Carpi, On synchronizing unambiguous automata, *Theoret. Comput. Sci.,* 60, 285–296, 1988.

[8] A. Carpi, F. D'Alessandro, The Synchronization Problem for Strongly Transitive Automata, in: M. Ito, M. Toyama eds., *DLT 2008 Developments in Language Theory, Lecture Notes in Computer Science*, Vol. 5257, pp. 240–251, Springer, Berlin, 2008.

[9] A. Carpi, F. D'Alessandro, Strongly transitive automata and the Černý conjecture *Acta Informatica,* 46, 591–607, 2009.

[10] A. Carpi, F. D'Alessandro, The synchronization problem for locally strongly transitive automata, in: R. Královič, D. Niwiński eds., *MFCS 2009 Mathematical Foundations of Computer Science, Lecture Notes in Comput. Sci.,* Vol. 5734, pp. 211–222, Springer, Berlin, 2009.

[11] A. Carpi, F. D'Alessandro, On the Hybrid Černý-Road coloring problem and Hamiltonian paths, in: Y. Gao, H. Lu, S. Seki, S. Yu eds., *DLT 2010 Developments in Language Theory, Lecture Notes in Comput. Sci.,* Vol. 6224, pp. 124–135, Springer, Berlin, 2010.

[12] A. Carpi, F. D'Alessandro, Independent sets of words and the synchronization problem, *Advances in Applied Mathematics,* 50, 339–355, 2013.

[13] A. Carpi, F. D'Alessandro, Černý-like problems for finite sets of words, Bistarelli S., Formisano A. eds., *15th Italian Conference on Theoretical Computer Science,* Vol. 1231, pp. 81–92, CEUR-WS.org 2014.

[14] A. Carpi, F. D'Alessandro, Completeness and synchronization for finite sets of words, *15th Mons Theoretical Computer Science Days,* Nancy (France) 2014.

[15] J. Černý, Poznámka k. homogénnym experimenton s konečnými automatmi, *Mat. fyz. cas SAV,* 14, 208–215, 1964.

[16] A. de Luca, F. D'Alessandro, Teoria degli Automi Finiti, 68, Springer Italia, 2013.

[17] G. Fici, E. V. Pribavkina, J. Sakarovitch, On the Minimal Uncompletable Word Problem, CoRR, arXiv: 1002.1928, 2010.

[18] V. V. Gusev, E. V. Pribavkina, On Non-complete Sets and Restivo's Conjecture, in: G. Mauri, A. Leporati eds., *DLT 2011 Developments in Language Theory, Lecture Notes in Comput. Sci.,* Vol. 6795, pp. 239–250, Springer, Berlin, 2011.

[19] J. E. Pin, *Le problème de la synchronization et la conjecture de Cerny,* Thèse de 3ème cycle, Université de Paris 6, 1978.

[20] J. E. Pin, Sur un cas particulier de la conjecture de Cerny, in: G. Ausiello, C. Böhm eds., *5th ICALP Lecture Notes in Computer Science*, Vol. 62, pp. 345–352, Springer, Berlin, 1978.

[21] A. Restivo, Some remarks on complete subsets of a free monoid, in: A. de Luca ed., *Non-Commutative Structures in Algebra and Geometric Combinatorics, International Colloquium, Arco Felice, July 1978, Quaderni de "La Ricerca Scientifica"*, CNR, 109, 19–25, 1981.

[22] A. N. Trahtman, The road coloring problem, *Israel J. Math.,* 172, 51–60, 2009.

[23] M. V. Volkov, Synchronizing automata and the Cerny conjecture, in: C. Martín-Vide, F. Otto, H. Fernau eds., *LATA 2008 Language and Automata Theory and Applications, Lecture Notes in Comput. Sci.,* Vol. 5196, pp. 11–27, Springer, Berlin, 2008.