COMPLETELY REACHABLE ALMOST GROUP AUTOMATA[1]

David Fernando Casas Torres

Ural Federal University,
51 Lenina Str., Ekaterimburg, 620000, Russian Federation

dkasastorres@urfu.ru

Abstract: We consider finite deterministic automata such that their alphabets consist of exactly one letter of defect 1 and a set of permutations of the state set. We study under which conditions such an automaton is completely reachable. We focus our attention on the case when the set of permutations generates a transitive imprimitive group.

Keywords: Deterministic finite automata, Transition monoid, Complete reachability, Permutation group.

1. Introduction

A deterministic finite automaton is said to be *completely reachable* if every non-empty subset of states is the image of the whole state set by the action of some word. Such automata appeared in the study of descriptional complexity of formal languages [11] and in relation to the Černý conjecture [8]. A systematic study of completely reachable automata was initiated in [5] and [6], and continued in [4]; in these papers Bondar and Volkov developed a characterization of completely reachable automata that relied on the construction of a series of digraphs.

One of the main results by Ferens and Szykuła [9] was an algorithm of polynomial time complexity, with respect to the number of states and letters, to decide whether a given automaton was completely reachable. This seemed to solve the complexity problem for this kind of automata. A different approach was proposed by Volkov and the author [7] for the special case of automata with two letters. There the characterization relied on whether one of the letters acted as a complete cyclic permutation of the states and how the other letter acted on certain subsets of states.

In this paper we give an approach to the generalization of the result in [7] by allowing that all the letters except one act as permutations of the set state and studying how the additional non-permutation letter acts on non-trivial blocks of imprimitivity if there are any.

The study of automata where all letters but one are permutations is by no means new. This kind of automata is presented with different names. In [1], they are called *almost-permutation automata* and are used to present an example of a series of slowly synchronzing automata with a sink state. In [2], automata are under the disguise of transformation semigroups and are called *near permutation*. In [3], they are called *almost-group automata*; there it is proved that these automata synchronize with high probability. Finally, in [12], the non-permutation letters are the identity except in a subset of states where they have the same image. There, it is proved that if no equivalence relation is preserved under the action of the letters, then the automaton is synchronizing. Among these papers, we would like to highlight the work done in [10] where the primitivity of a group of permutations of a state set has been tightly related to the complete reachability of the automata

---

generated by adding a non permutation letter. Thus, the result presented here approaches this theory from the other side where the group is transitive but not primitive and we suggest a condition to ensure that the automaton generated is completely reachable.

In Section 2 we present the definitions and notation used in this paper. Then in Section 3 we present and prove a necessary condition for almost group automata to be completely reachable. In Section 4 we describe a set of directed graphs useful in the discussion of complete reachability and prove a partial sufficient condition related with the one discussed in the Section 3.

2. Preliminaries

A *deterministic finite automaton*, or simply an automaton, is usually defined as a triple

$$\mathcal{A} = \langle Q, \Sigma, \delta \rangle,$$

where Q, the states, and Σ, the alphabet, are finite sets and $\delta : Q \times \Sigma \to Q$ is the transition function. For each letter in alphabet of the automaton $a \in \Sigma$ we can define the function $\delta_a : Q \to Q$ where $\delta_a(q) = \delta(q, a)$, hence each letter can be considered individually as a function of Q on itself or a *transformation* over Q. This observation makes reasonable for us to use the following notation: for every $q \in Q$ and $a \in \Sigma$ we will denote $\delta(q, a) := q \cdot a$. Derived from this we can say that an automaton can be specified using just its set of states and the action of each letter in this set; that is why from now on we will define automata as pairs of the state set and the alphabet.

A *word* of an automaton is a finite sequence of letters over its alphabet; this includes the empty word. The set of all words over the alphabet Σ is denoted by Σ^*. We can extend the action of letters to words recursively in the following way: if $w \in \Sigma^*$, $a \in \Sigma$ and $q \in Q$, then $q \cdot wa := (q \cdot w) \cdot a$, and the action of the empty word is the identity function. Furthermore, the action of words can be applied to subsets of states: if $P \subseteq Q$ and $w \in \Sigma^*$, then

$$P \cdot w := \{p \cdot w \mid \text{ for every } p \in P\}.$$

A subset of states $P \subseteq Q$ is called *reachable* if there is a word $w \in \Sigma^*$ such that its image is exactly P, that is, $Q \cdot w = P$. An automaton is said to be *synchronizing* if at least one singleton is reachable, i.e., there is a state $q \in Q$ and a word $w \in \Sigma^*$ such that $Q \cdot w = \{q\}$. An automaton is *completely reachable* if every non-empty subset of states is reachable.

Let $\mathcal{A} = \langle Q, \Sigma \rangle$ be an automaton and $w \in \Sigma^*$ an arbitrary word. The excluded set of w denoted by $\mathrm{excl}(w)$ is the set of states that have no preimages by w. The *defect* of w is the size of its excluded set. In the case the defect is 0, the word w represents a permutation of the set of states. Since a word is a total function, if the defect of w is bigger than 0, then there must be states with the same image. These images are the *duplicated* states of the word; the set of these states will be denoted by $\mathrm{dupl}(w)$. When any of these sets, $\mathrm{excl}()$ or $\mathrm{dupl}()$, is a singleton we will make no distinction between the set and the state inside it. Additionally, for the case of words of defect 1 we know that exactly two states must have the same image; we will call this pair of states the *collapsed set* of the word, denoted by $\mathrm{coll}()$.

Some transformations over a set of states Q can be bijective, thus permutations. Thanks to this we can use some terminology of the theory of permutation groups. Recall that the set of all the bijective transformations of a finite set Q on itself is denoted by S_Q, also called the *symmetric group*. Let $G \subseteq S_Q$ be a group of permutations of Q. This group is said to be *transitive* if for every pair of states $q, p \in Q$ there is a permutation $\sigma \in G$ such that $p \cdot \sigma = q$. Our main subject of study are automata and their words, thus, except when specified, when we talk about arbitrary permutations of the set of states we assume that there is allways a word that produces it. This is,

if we mention the permutation $g \in S_Q$, we assume that there is a word w such that $q \cdot w = q \cdot g$ for every $q \in Q$.

A non-empty subset $B \subseteq Q$ is said to be a *block* of the group if and only if for every $\sigma \in G$ either $B \cdot \sigma = B$ or $B \cdot \sigma \cap B = \emptyset$. The singletons and Q itself are, always, blocks, these are called *trivial*. A permutation group $G \subseteq S_Q$ is said to be *primitive* if it is transitive and the only blocks are the trivial ones; otherwise the group is said to be *imprimitive*. In this article when we talk about a block of imprimitivity, unless stated the contrary, it will always be non-trivial. If a transitive group $G \subset S_Q$ has a block of imprimitivity $B \subseteq Q$, all the images of B by elements of G are also blocks of imprimitivity and form a partition of Q. This partition of subsets is called a *system of imprimitivity* of the group G over Q.

A *directed graph* Γ is a pair $(V(\Gamma), E(\Gamma))$, where $V(\Gamma)$ is the vertex set and $E(\Gamma) \subseteq V(\Gamma) \times V(\Gamma)$ is the set of directed edges. To each edge we can assign one or more labels from some set L. We will denote an edge $(s, t) \in E(\Gamma)$ labelled with w as $s \xrightarrow{w} t$. Since in this paper we consider only directed graphs, from now we omit the word "directed". For reference, the first and second components of an edge will be called the *source* and *tail* respectively. A *path* of a graph is a set of edges $e_1, e_2, \ldots, e_m$, $m \geq 1$, such that for every $1 \leq i < m$, the tail of e_i is the same as the source of e_{i+1}. Vertices $p, q \in V(\Gamma)$ are *strongly connected* if there is a path from p to q and from q to p. Also we consider each vertex as strongly connected with itself. A strongly connected component of a graph is a maximal subset of vertices such that all its vertices are strongly connected to each other.

An automaton can be represented as a labelled graph, where the vertex set is the states set and for each state p and letter a, there is a labelled edge $p \xrightarrow{a} p \cdot a$. This is the *underlying* graph of the automaton.

As we mentioned in the introduction, in [10] the following characterization of primitive permutation groups is given. Here $[n] := \{1, 2 \ldots, n\}$ and if S is a set of transformations, then $\langle S \rangle$ is the transformation semigroup generated by S. Recall that a transformation f is *idempotent* if $f^2 = f$.

Theorem 1 [10, Theorem 3.1]. *Let G be a permutation group on $[n]$ with $n \geq 3$. Then G is primitive if and only if for each [idempotent] transformation $f : [n] \to [n]$ of defect 1 every non-empty subset $A \subseteq [n]$ is reachable in $\langle G \cup \{f\} \rangle$.*

This theorem presents a characterization of primitive groups. In the language of automata it states that in the presence of a set of permutation letters that generates a primitive group, the addition of any transformation of defect 1 suffices to obtain a completely reachable automaton. Here we study a related case. We would like to know what happens when the group generated by the permutation letters is transitive but not primitive. We will see how this situation is not that forgiving and it requires a more complex relation between the group generated by the permutations and the transformation of defect 1. The results proved in this article are closely related to the ones presented in [7] for automata with just one permutation letter and one with defect 1. The work presented in this article and in [7] is based on the work made in [6]; there the main actor is a graph constructed in several steps. We will briefly explain the construction in Section 4.

For the rest of this paper we will consider automata $\mathcal{A} = \langle Q, \Sigma \rangle$, where $\Sigma = \Sigma_0 \cup \{a\}$ and:

- the set of letters $\Sigma_0 \subset S_Q$, are all permutations of Q,
- the generated subgroup $G = \langle \Sigma_0 \rangle$ is transitive,
- the letter a has defect 1.

The excluded state of the only letter of defect 1 will be denoted by e, i.e., $\mathrm{excl}(a) = e$. Unless specified otherwise, the group generated by all permutation letters is denoted as G. We will call automata with these characteristics *almost group* automata.

Let $r \in \mathrm{coll}(a)$ be one of the two states collapsed by a. There is a permutation that sends e to r; call it $\sigma \in G$. The transformation σa has defect 1, $e = \mathrm{excl}(\sigma a)$, and $e \in \mathrm{coll}(\sigma a)$. Consider the

automaton $\overline{\mathcal{A}} = \langle Q, \Sigma_0 \cup \{\sigma a\} \rangle$. Note that $\mathcal{A}$ is completely reachable if and only if $\overline{\mathcal{A}}$ is completely reachable. Therefore, there is no loss of generality when we add the condition that $e \in \mathrm{coll}(a)$ from the beginning. When this happens we call the automaton *standardized*. This change will simplify the arguments we use for the rest of the article.

For any automaton a subset of states, $P \subseteq Q$ is *invariant* by a transformation $w \in \Sigma^*$, or *w-invariant*, if $P \cdot w \subseteq P$. The condition to get complete reachability in the binary case is that no subset of states that *represents* a subgroup of the cyclic group is invariant by letter of defect 1. The cosets of any subgroup generate a partition of the group that contains this subgroup. There is a parallel situation in the case of blocks of imprimitivity, they form a partition of the set of states. There are subgroups for each of these blocks of imprimitivity that let them invariant[2]. This is the main reason to consider systems of imprimitivity. These partitions of the set of states are the closest to represent subgroups of the group acting on the states.

3. A necessary condition

First, we begin by proving that complete reachability implies that some blocks can not be a-invariant.

Proposition 1. *Let $\mathcal{A}$ be a standardized almost group automaton. If $\mathcal{A}$ is completely reachable, then G is transitive and if there is at least one block of imprimitivity then no block of imprimitivity that contains $e = \mathrm{excl}(a)$ is invariant by a.*

P r o o f. First, let us prove that the condition for the group generated by the set of permutations to be transitive is necessary. For every word $w \in \Sigma^*$ it is true that $e \in \mathrm{excl}(wa)$; furthermore,

$$|\mathrm{excl}(w)| \leq |\mathrm{excl}(wa)| \leq |\mathrm{excl}(w)| + 1.$$

This is, the action of adding a to a word either increases by one or keeps the defect of the resulting word. Note that adding a permutation does not modify the defect of any word. Hence, in order to reach the subsets $Q \setminus \{q\}$ for every $q \in Q$, it is necessary that there exists a permutation $\sigma_q \in G$ such that $e \cdot \sigma_q = q$. Let $p, q \in Q$ be an arbitrary pair of states. By the previously said, if $\mathcal{A}$ is completely reachable, then there are two permutations $\sigma_p, \sigma_q \in G$ such that $e \cdot \sigma_p = p$ and $e \cdot \sigma_q = q$. Finally, note that $p \cdot \sigma_p^{-1} \sigma_q = q$. Thus, G is transitive.

For a subset of states $S \subset Q$, we denote by $\overline{S}$ its complement, i.e., $Q \setminus S$.

The proof that no block of imprimitivity is a-invariant will be by contradiction. Suppose that B is a block of imprimitivity that contains e and is a-invariant. This block belongs to a system of imprimitivity. Let $w \in \Sigma^*$ be the shortest word that reaches the complement of a block of this system, say $\overline{C}$. If $w = w' b$ with $b \in \Sigma_0$, then $Q \cdot w' = \overline{C} \cdot b^{-1}$, the complement of a block of imprimitivity. This contradicts the condition of w being the shortest word. Hence, w can not end in a permutation.

As a consequence, the word w ends with the letter a, i.e., $w = w' a$. Recall that $Q \cdot w' a \subset Q \cdot a$ and $e \notin Q \cdot a$, thus $e \notin Q \cdot w$ and we conclude that $Q \cdot w = \overline{B}$.

Since B is a-invariant, we can conclude that its complement is also a-invariant. And since every $q \in \overline{B}$ has a preimage by a then this letter acts as a permutation of $\overline{B}$. Therefore $Q \cdot w' = Q \cdot w' a = \overline{B}$ what, again, contradicts the supposition of w being the shortest. There is no other type of letter in which the word w could finish, then we end with an absurd. This situation came from supposing that B is a-invariant, thus we have our proposition.   □

By the preceding proof we have:

---

[2]Considerations of this are treated ahead in the paper.

Corollary 1. *If there is a block of imprimitivity that contains e and is invariant by a, then its complement is not reachable.*

4. Rystsov graphs of almost group automata

4.1. The structure of Rystsov graphs

In [6] and [4] Bondar and Volkov presented a characterization of completely reachable automata. The characterization relies on the construction of a graph that can be constructed from an arbitrary automaton. This graph is a generalization of the ideas presented by Igor Rystsov in [13]. That is why we will call these graphs as *Rystsov graphs* of the automata.

The Rystsov graph of an automaton $\mathcal{A}$, denoted by $\Gamma(\mathcal{A})$, is constructed in an inductive way. This means that in order to construct $\Gamma(\mathcal{A})$ we first assemble a graph called $\Gamma_1(\mathcal{A})$, verify if we can continue and if that is the case from $\Gamma_1(\mathcal{A})$ we continue with the construction of $\Gamma_2(\mathcal{A})$ and so on. This series of graphs is guaranteed to always finish, the final graph is the Rystsov graph of $\mathcal{A}$. For the construction of the partial graphs we make use of the sets of words $W_k(\mathcal{A}) \subset \Sigma^*$, defined as the subset of all words of defect k for $k \geq 1$.

The first step is to construct the graph $\Gamma_1(\mathcal{A})$ where its vertex set is $Q_1 := Q$ and its edge set is defined as:

$$E_1(\mathcal{A}) := \{\operatorname{excl}(w) \xrightarrow{w} \operatorname{dupl}(w) \in Q_1 \times Q_1 \mid w \in W_1(\mathcal{A})\}.$$

Example 1. Consider the automaton $\mathcal{E}_{18} := \langle \{1, 2, \ldots, 18\}, \{a, b, c\}\rangle$, where b and c are permutations with the following cyclic representation:

$$
\begin{aligned}
b :=& (1, 11, 13, 5, 7, 17)(2, 10, 14, 4, 8, 16)(3, 12, 15, 6, 9, 18), \\
c :=& (1, 3, 2)(4, 5, 6)(7, 13)(8, 16)(9, 15)(10, 14)(11, 17)(12, 18),
\end{aligned}
$$

and the transformation a has defect 1. The following representation of a puts the respective image under each state and omits the states that do not change:

$$
\begin{pmatrix} 1\ 2\ 5\ 6\ 8 \\ 6\ 8\ 6\ 5\ 2 \end{pmatrix}.
$$

Note that the $\operatorname{excl}(a) = 1$, $\operatorname{dupl}(a) = 6$ and $\operatorname{coll}(a) = \{1, 5\}$. The group generated by $\{b, c\}$ is transitive and the blocks of imprimitivity that contain the state 1 are the sets

$$\{1, 5\}, \quad \{1, 2, 3, 4, 5, 6\}.$$

We can see $1 \xrightarrow{a} 6$, $2 \xrightarrow{ac^2} 5$ and $1 \xrightarrow{ab^3a} 3$ are edges in $E_1(\mathcal{E}_{18})$.

We continue the definitions. For any automaton $\mathcal{A}$ consider the following subset of states:

$$D_1(\mathcal{A}) := \{p \in Q_1 \mid p = \operatorname{dupl}(w) \quad \& \quad e = \operatorname{excl}(w) \text{ for } w \in W_1(\mathcal{A})\}.$$

These are the states directly connected to e in $\Gamma_1(\mathcal{A})$, that is, tails of the edges with e as source. The following lemma states that all the edges in $\Gamma_1(\mathcal{A})$ are images by G of these initial edges.

Lemma 1. *If $q \to p \in E(\Gamma_1(\mathcal{A}))$, then there are $\sigma_q \in G$ and $d \in D_1(\mathcal{A})$ such that $e \cdot \sigma_q = q$ and $d \cdot \sigma_q = p$. Or, what is equivalent, there is a permutation $\sigma_q \in G$ such that $p \cdot \sigma_q^{-1} \in D_1(\mathcal{A})$.*

P r o o f. If $q \to p$ is an edge of $\Gamma_1(\mathcal{A})$, then there is a word of defect 1, call it $w \in W_1(\mathcal{A})$ such that $\mathrm{excl}(w) = q$ and $\mathrm{dupl}(w) = p$; this happens due to the definition of $\Gamma_1(\mathcal{A})$. Remember that G is transitive, thus there is a permutation $\sigma_q \in G$ such that $e \cdot \sigma_q = q$.

The word $w\sigma_q^{-1}$ has defect 1 and $\mathrm{excl}(w\sigma_q^{-1}) = \mathrm{excl}(w) \cdot \sigma_q^{-1}$, at the same time $\mathrm{dupl}(w\sigma_q^{-1}) = \mathrm{dupl}(w) \cdot \sigma_q^{-1}$. Thus $\mathrm{excl}(w\sigma_q^{-1}) = q \cdot \sigma_q^{-1} = e$ and $p \cdot \sigma_q^{-1} \in D_1(\mathcal{A})$. $\qquad\square$

This lemma also tell us that in order to compute $\Gamma_1(\mathcal{A})$ it is sufficient to calculate $D_1(\mathcal{A})$, and then apply to the generated edges permutations that send e to each of the different states of the automaton. In our running example the initial edges of $\Gamma_1(\mathcal{E}_{18})$ are shown in Figure 1. The
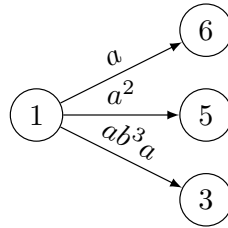

Figure 1. The initial edges of $\Gamma_1(\mathcal{E}_{18})$.

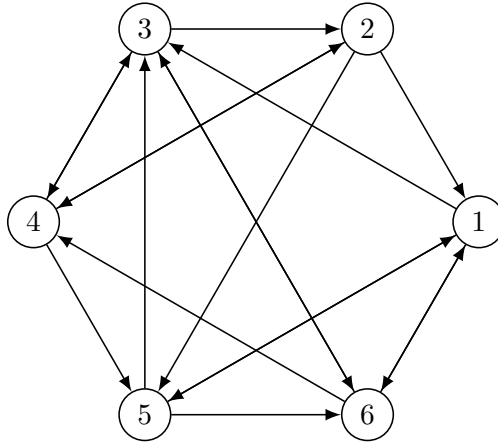strongly connected component that contains 1 is shown in Figure 2 (we omitted the labels to avoid confusion).

Figure 2. A strongly connected component of $\Gamma_1(\mathcal{E}_{18})$.

Now, let $C_e^{[1]} \subseteq Q_1$ be the vertex set of the strongly connected component of $\Gamma_1(\mathcal{A})$ that contains e.

Lemma 2. *The set $C_e^{[1]}$ is a block of imprimitivity.*

P r o o f. If $\Gamma_1(\mathcal{A})$ is strongly connected then $C_e^{[1]} = Q$ and the proposition is true. Then let us assume $\Gamma_1(\mathcal{A})$ is not strongly connected and $C_e^{[1]}$ is a proper subset of Q.

Let $\sigma \in G$ be a permutation such that $e \cdot \sigma = \mathrm{dupl}(a) = d$, then the edge

$$d \xrightarrow{a\sigma} d \cdot \sigma \in E_1.$$

If we repeat the application of σ, the produced words have all defect 1, and there is an $i \geq 1$ such that $d \cdot \sigma^i = e$. Then, $C_e^{[1]}$ is not a singleton since at least $e, d \in C_e^{[1]}$.

Considering this we will prove first that for any $\rho \in G$, the subset $C_e^{[1]} \cdot \rho$ is also a strongly connected component.

Let $p, q \in C_e^{[1]}$ be two arbitrary states. In $\Gamma_1(\mathcal{A})$ there is a path:

$$p \xrightarrow{w_1} t_1 \xrightarrow{w_2} t_2 \ldots t_{k-1} \xrightarrow{w_k} q$$

where every w_i is a word of defect 1. Since permutations act well on excluded and duplicated states, then:

$$p \cdot \rho \xrightarrow{w_1\rho} t_1 \cdot \rho \xrightarrow{w_2\rho} t_2 \cdot \rho \ldots t_{k-1} \cdot \rho \xrightarrow{w_k\rho} q \cdot \rho$$

is a path in $C_e^{[1]} \cdot \rho$; in the same way we can prove the existence of a path connecting $q \cdot \rho$ with $p \cdot \rho$, making $C_e^{[1]} \cdot \rho$ a strongly connected component.

What is left is to prove that $C_e^{[1]}$ and its images by permutations of G are blocks of imprimitivity. Let $\rho \in G$ be an arbitrary permutation. Suppose $x \in C_e^{[1]} \cap C_e^{[1]} \cdot \rho$ and let $y \in C_e^{[1]}$ and $z \in C_e^{[1]} \cdot \rho$ be two different states. By the definition of strongly connected component there are paths from y to x, from x to z and, going back, from z to x, and from x to y. Then

$$C_e^{[1]} = C_e^{[1]} \cdot \rho.$$

This makes $C_e^{[1]}$ a block of imprimitivity. $\qquad\square$

We continue the inductive construction of the graph $\Gamma(\mathcal{A})$. Once we get $\Gamma_k(\mathcal{A})$, $k \geq 1$, if one of the following alternatives happens then the construction will be stopped and $\Gamma(\mathcal{A}) := \Gamma_k(\mathcal{A})$: either the graph is strongly connected, or all the strongly connected components are not *big enough* (we will address the meaning of this in a moment). If none of these two possibilities happen, then we proceed to construct $\Gamma_{k+1}(\mathcal{A})$. The new vertex set Q_{k+1} will consists of the strongly connected components of $\Gamma_k(\mathcal{A})$; thus, each vertex is a collection of vertices of the set Q_k.

In order to define the edges of this new graph, we need to properly define when a strongly connected component is *big enough*. For this note that each vertex of $\Gamma_2(\mathcal{A})$ is a subset of states (even considering singletons) and the vertices of $\Gamma_3(\mathcal{A})$ would be collections of subsets of states and so on. With this in mind for $k \geq 2$ let $V \in Q_k$ be a vertex of $\Gamma_k(\mathcal{A})$, define the *foliage* of V, and denote it by $\mathrm{leaf}(V)$, as follows: for $V \in Q_2$, its foliage is the set itself, i.e., $\mathrm{leaf}(V) := V$, and for $k > 2$,

$$\mathrm{leaf}(V) := \bigcup_{x \in V} \mathrm{leaf}(x).$$

At the end $\mathrm{leaf}(V)$ is a subset of states. A vertex V of $\Gamma_{k+1}(\mathcal{A})$, or, what is the same, a strongly connected component of $\Gamma_k(\mathcal{A})$, is *big enough* if $|\mathrm{leaf}(V)| \geq k + 1$. Thus, we stop the construction if none of the would be vertices of $\Gamma_{k+1}(\mathcal{A})$ have more of $k + 1$ states in their foliages. (The term "foliage" is borrowed from [4], where the definition of the vertex sets of the graphs Γ takes form of a rooted tree.) Suppose that this is not the case and we can continue the process, then we can define a new set of edges:

$$E_{k+1} := \{C \xrightarrow{w} D \in Q_{k+1} \times Q_{k+1} \mid C \neq D, \text{ there is a } w \in W_{k+1}(\mathcal{A}),$$
$$\mathrm{excl}(w) \subseteq \mathrm{leaf}(C), \ \mathrm{dupl}(w) \cap \mathrm{leaf}(D) \neq \emptyset\}.$$

The edge set of $\Gamma_{k+1}(\mathcal{A})$ will be the edges of $\Gamma_k(\mathcal{A})$ that connect different vertices of Q_{k+1} together with the set E_{k+1}. For a more detailed discussion of the construction of $\Gamma(\mathcal{A})$ we recommend the reader [4, Section 3].

We have the following theorem that characterizes completely reachable automata:

Theorem 2. *[6] If an automaton $\mathcal{A} = \langle Q, \Sigma \rangle$ is such that the graph $\Gamma(\mathcal{A})$ is strongly connected and $\Gamma(\mathcal{A}) = \Gamma_k(\mathcal{A})$, then $\mathcal{A}$ is completely reachable; more precisely, for every non-empty subset $P \subseteq Q$, there is a product w of words of defect at most k such that $P = Q \cdot w$.*

In the case that the group G is primitive over Q, from Lemma 2 we can see that $\Gamma_1(\mathcal{A})$ will be strongly connected and by Theorem 2 it immediately follows that $\mathcal{A}$ is completely reachable. That is why, from now on the group G will, besides being transitive, have at least a block of imprimitivity.

Example 2.   Recall the automaton $\mathcal{E}_{18}$ presented in Example 1. We have seen that $C_e^{[1]} = \{1, 2, 3, 4, 5, 6\}$, and the other strongly connected components are the sets $B_2 = \{7, 8, 9, 10, 11, 12\}$ and $B_3 = \{13, 14, 15, 16, 17, 18\}$. Since these sets have more than two elements, we can continue the construction of $\Gamma(\mathcal{E}_{18})$. Accordingly to the previously said, the vertex set of $\Gamma_2(\mathcal{E}_{18})$ is $Q_2 = \{C_e^{[1]}, B_2, B_3\}$. Consider the word $w := ab^3aca$, note that $\mathrm{excl}(w) = \{1, 3\}$ and $\mathrm{dupl}(w) = \{8, 6\}$, hence the edge $C_e^{[1]} \xrightarrow{w} B_2 \in E_2$. If we add b twice more we have:

$$\mathrm{excl}(wb) = \{11, 12\}, \quad \mathrm{dupl}(wb) = \{9, 16\}$$
$$\mathrm{excl}(wbb) = \{13, 15\}, \quad \mathrm{dupl}(wbb) = \{18, 2\}.$$

Thus adding the edges $B_2 \xrightarrow{wb} B_3$ and $B_3 \xrightarrow{wbb} C_e^{[1]}$ to E_2. These are enough to conclude, thanks to Theorem 2, that $\mathcal{E}_{18}$ is completely reachable.

We will extend the results given by Lemma 1 and Lemma 2. Following the previous notation denote the strongly connected component that contains e in the graph $\Gamma_k(\mathcal{A})$ as $C_e^{[k]}$.

Lemma 3. *If the foliages of the vertices in $\Gamma_k(\mathcal{A})$ form a system of imprimitivity of G over Q, then $Y \to Z \in E_{k+1}$ if and only if there is a permutation $\sigma \in G$ and a set $X \in Q_k$ such that $\mathrm{leaf}(Y) = \mathrm{leaf}(C_e^{[k]}) \cdot \sigma$; $\mathrm{leaf}(X) \cdot \sigma = \mathrm{leaf}(Z)$ and $C_e^{[k]} \to X \in E_{k+1}$.*

P r o o f.   Since permutations respect the defect of any word and act well on excluded and duplicated sets, the converse is easy to see.

Now, if $Y \xrightarrow{w} Z \in E_{k+1}$, with $w \in W_{k+1}(\mathcal{A})$, then $\mathrm{excl}(w) \subset \mathrm{leaf}(Y)$ and $\mathrm{dupl}(w) \cap \mathrm{leaf}(Z) \neq \emptyset$. Let $w = u\, a\, \sigma$ with $\sigma \in \Sigma^*$ as a permutation, this is, σ is the longest word generating a permutation after the last appearance of the letter a in w. Since permutations do not increase the defect of a word, then $ua \in W_{k+1}(\mathcal{A})$ and $\mathrm{excl}(w\sigma^{-1}) = \mathrm{excl}(ua)$. From the last affirmation we can conclude that $\mathrm{excl}(ua) \subseteq \mathrm{leaf}(Y) \cdot \sigma^{-1}$.

Since, by hypothesis, $\mathrm{leaf}(Y)$ is a block of imprimitivity then also it is $\mathrm{leaf}(Y) \cdot \sigma^{-1}$. Recall that $e \in \mathrm{excl}(ua)$ thus $\mathrm{excl}(ua) \subseteq C_e^{[k]} = \mathrm{leaf}(Y) \cdot \sigma^{-1}$. Using the same argument we can conclude that $\mathrm{leaf}(X) = \mathrm{leaf}(Z) \cdot \sigma^{-1}$.   □

Lemma 4. *If the foliages of the vertices in $\Gamma_k(\mathcal{A})$ form a system of imprimitivity, then the foliage of $C_e^{[k+1]}$ is a block of imprimitivity of G over Q.*

P r o o f.   If each of the foliages of the vertices of $\Gamma_k(\mathcal{A})$ forms a system of imprimitivity, then the foliage of $C_e^{[k+1]}$ is just the union of blocks of imprimitivity.
We can use an argument similar to the one used in the proof of Lemma 2 to prove that the image by any $\sigma \in G$ of the foliage of $C_e^{[k+1]}$ is also a strongly connected component and a block of imprimitivity.   □

Lemma 3 and Lemma 4 form the proof by induction the following result.

Proposition 2. *For any $k \geq 1$, the foliages of the vertices of each $\Gamma_k(\mathcal{A})$ form a system of imprimitivity.*

Note that for any $k \geq 1$ if it happens that $\mathrm{leaf}(C_e^{[k]}) = Q$ then $\Gamma_k(\mathcal{A})$ is strongly connected and $\mathcal{A}$ is completely reachable. Now we will prove that if this is not the case for any k, then some block of imprimitivity that contains e is invariant by a. We will use the following set:

$$D_k(\mathcal{A}) := \{p \in Q \mid p \in \mathrm{dupl}(w) \text{ for some } w \in \Sigma^*$$
$$\text{such that } |\mathrm{excl}(w)| \leq k \text{ and } e \in \mathrm{excl}(w) \subseteq \mathrm{leaf}(C_e^{[k-1]})\}.$$

The set of states duplicated by words of defect less than k such that their excluded set is contained in $C_e^{[k]}$.

4.2. Intermezzo

Before we continue, it is necessary to present some definitions and results related to the theory of permutation groups that are used in the rest of this section. Let Q be a finite set and $G \subseteq S_Q$ be a subgroup of permutations of Q. For any non-empty subset $P \subset Q$ consider the set of permutations:

$$\mathrm{St}_G(P) := \{\sigma \in G \mid P \cdot \sigma = P\},$$

that is, the set of permutations of G that preserve P set-wise. It can be easily proved that $\mathrm{St}_G(P)$ is a subgroup of G. Let us call it the *stabilizer* of the subset P.

Now consider an arbitrary but fixed system of imprimitivity of G over Q, call it $\mathfrak{B}$. The following fact is well known and we will omit the proof.

Proposition 3. *Let G be a group of permutations of a finite set Q. Suppose that G is transitive and $\mathfrak{B}$ is a system of imprimitivity. If $B, C \in \mathfrak{B}$ are two different blocks of imprimitivity then $\mathrm{St}_G(B)$ and $\mathrm{St}_G(C)$ are conjugate subgroups of G.*

For a subgroup H of a group G, the *core* of H, denoted by $\mathrm{Cr}(H)$, is the intersection of all the conjugates of H in G, i.e.,

$$\mathrm{Cr}(H) := \bigcap_{\sigma \in G} \sigma^{-1} H \sigma.$$

Note that this subgroup is normal for G.

Resuming with the transitive group G of permutations of Q, Proposition 3 tells us that for every system of imprimitivity $\mathfrak{B}$ of Q all the stabilizers of the blocks are conjugate. Hence, the following definition makes sense.

Definition 1. *Let G be a subgroup of permutations of Q and $\mathfrak{B}$ be a system of imprimitivity of Q. The core of $\mathfrak{B}$, denoted by $\mathrm{Cr}(\mathfrak{B})$, is the intersection of all the stabilizers of the blocks in $\mathfrak{B}$.*

In some occasions it is more convenient to work with blocks of imprimitivity, hence to talk about the core of a block of imprimitivity. If $\mathfrak{B}$ is a system of imprimitivity and $B \in \mathfrak{B}$ is a block, we denote $\mathrm{Cr}(B) := \mathrm{Cr}(\mathfrak{B})$. For our purposes we look for the core of certain blocks of imprimitivity to act in a transitive way on said blocks. We can ensure this if said core acts transitively on at least one of the blocks.

Proposition 4. *Let G be a group of permutations of the finite set Q. Suppose that G is transitive and $\mathfrak{B}$ is a system of imprimitivity. If $B \in \mathfrak{B}$ is a block and $\mathrm{Cr}(\mathfrak{B})$ acts transitively on B, then this core is also transitive on all the blocks of $\mathfrak{B}$.*

P r o o f. Let $C \in \mathfrak{B}$ be a different block of $\mathfrak{B}$, besides let $p, q \in C$ be two different states. We aim to prove that there is a permutation $\sigma \in \text{Cr}(\mathfrak{B})$ such that $p \cdot \sigma = q$. Being G transitive, there is a permutation $\tau \in G$ such that $C \cdot \tau = B$. Let $r, s \in B$ be such that $p \cdot \tau = r$ and $q \cdot \tau = s$. By hypothesis, there is a permutation $\rho \in \text{Cr}(\mathfrak{B})$ such that $r \cdot \rho = s$. Thus

$$p \cdot \tau \rho \tau^{-1} = q.$$

Since the core is normal in G we can conclude that $\tau \rho \tau^{-1} \in \text{Cr}(\mathfrak{B})$.   □

4.3. Non-reachability and invariance

In this part we see that for some almost-group automata not being completely reachable implies there is at least one block of imprimitivity invariant by the letter of defect 1.

Before the main proposition we present a technical lemma. Since in the following lemma k is arbitrary but fixed, $C_e^{[k]}$ will be referred just as C_e.

Lemma 5. *Let $\mathcal{A} = \langle Q, \Sigma_0 \cup \{a\} \rangle$ be an almost-group automaton. If in $\Gamma_k(\mathcal{A})$ there is an edge $C_e \to X$ and $\text{Cr}(\text{leaf}(C_e))$ is transitive for C_e, then for every state $q \in \text{leaf}(X)$, there exists a word v of defect k such that $\text{excl}(v) \subset \text{leaf}(C_e)$ and $q \in \text{dupl}(v)$.*

P r o o f. The edge $C_e \to X$ is produced by a word w such that $\text{excl}(w) \subset \text{leaf}(C_e)$ and $\text{dupl}(w) \cap \text{leaf}(X) \neq \emptyset$. Let $p \in \text{dupl}(w) \cap \text{leaf}(X)$ be arbitrary. Since $\text{Cr}(C_e)$ is transitive, by Proposition 4 there is a permutation $\sigma \in \text{Cr}(C_e)$ such that $p \cdot \sigma = q$. At the same time it is true that $C_e \cdot \sigma = C_e$, since the core is a subset of $\text{St}_G(C_e)$. Therefore we have that $\text{excl}(w\sigma) \subset \text{leaf}(C_e)$ and $q \in \text{dupl}(w\sigma)$.   □

Using the Lemma 3 we also can conclude:

Corollary 2. *If in $\Gamma_k(\mathcal{A})$ there is an edge $X \xrightarrow{w} Y$ and $\text{Cr}(\text{leaf}(C_e))$ is transitive for C_e. Then for every state $q \in \text{leaf}(Y)$, there exists a word v of defect k such that $\text{excl}(v) \subseteq \text{leaf}(X)$ and $q \in \text{dupl}(v)$.*

With these two lemmas, we are ready for the main result of this part:

Theorem 3. *Let $\mathcal{A} = \langle Q, \Sigma_0 \cup \{a\} \rangle$ be an almost-group automaton. Suppose $\Gamma(\mathcal{A})$ is not strongly connected. This means for some $k \geq 1$ it happens that $\Gamma(\mathcal{A}) = \Gamma_k(\mathcal{A})$; and $C_e^{[k]} = C_e^{[j]}$ for every $j \geq k$. Besides this, suppose that for every $\ell \leq k$ the cores $\text{Cr}(C_e^{[\ell]})$ are transitive on $C_e^{[\ell]}$. Then $\text{leaf}(C_e^{[k]})$ is invariant by a.*

P r o o f. We will use a, structurally, similar proof of the same fact for binary automata presented in [7]. Suppose that $C_e^{[k]} = C_e^{[k+1]}$. By induction on $0 \leq \ell \leq k$ we will prove that

$$\text{leaf}(C_e^{[\ell]}) \cdot a \subseteq \text{leaf}(C_e^{[k]}).$$

For $\ell = 0$ take $C_e^{[0]} = \{e\}$ hence the proposition is true in this case.

Our *first induction hypothesis* is that $\text{leaf}(C_e^{[\ell]}) \cdot a \subseteq \text{leaf}(C_e^{[k]})$. By the construction of $\Gamma_{\ell+1}(\mathcal{A})$, for any $p \in \text{leaf}(C_e^{[\ell+1]})$ there is a $X_m \in Q_\ell$ such that $p \in \text{leaf}(X_m)$ and there is a path:

$$C_e^{[\ell]} \to X_1 \to X_2 \to \cdots \to X_m$$

in $\Gamma_\ell(\mathcal{A})$.

Now, by induction on the length of the path (the number $m > 1$) the idea is to prove that $\mathrm{leaf}(X_m) \cdot a \subseteq \mathrm{leaf}(C_e^{[k]})$.

If $m = 1$, since there is an edge $C_e^{[\ell]} \to X_1$ we use Lemma 5 to ensure that for $p \in \mathrm{leaf}(X_1)$ there is a word $w \in W_\ell(\mathcal{A})$ such that $\mathrm{excl}(w) \subseteq \mathrm{leaf}(C_e^{[\ell]})$ and $p \in \mathrm{dupl}(w) \cap \mathrm{leaf}(X_1)$. The defect of wa is at most $\ell + 1 \leq k + 1$ and by the *first induction hypothesis* $\mathrm{excl}(wa) \subseteq \mathrm{leaf}(C_e^{[k]})$ and

$$p \cdot a \in \mathrm{dupl}(wa) \subseteq D_{k+1}(\mathcal{A}) \subseteq C_e^{[k+1]} = C_e^{[k]},$$

proving what we wanted.

Now suppose that $m > 1$ and $\mathrm{leaf}(X_{m-1}) \cdot a \subseteq \mathrm{leaf}(C_e^{[k]})$, i.e., the *second induction hypothesis*. By the Corollary 2 for $p \in \mathrm{leaf}(X_m)$ there is a word $w \in W_\ell(\mathcal{A})$ such that $\mathrm{excl}(w) \subseteq \mathrm{leaf}(X_{m-1})$ and $p \in \mathrm{dupl}(w)$. If we apply the same argument as before, but this time using the *second induction hypothesis* we can conclude that

$$p \cdot a \in \mathrm{dupl}(wa) \subseteq D_{k+1}(\mathcal{A}) \subseteq C_e^{[k+1]} = C_e^{[k]},$$

again, as intended.

Since $C_e^{[\ell+1]}$ is a strongly connected component of $\Gamma_\ell(\mathcal{A})$, thus its foliage is the union of the respective foliages of its vertices. We have proved that

$$\mathrm{leaf}(C_e^{[\ell+1]}) \cdot a \subseteq \mathrm{leaf}(C_e^{[k+1]}) = \mathrm{leaf}(C_e^{[k]}).$$

$\square$

The previous theorem proves that for certain almost group automata not being completely reachable is equivalent to having a non-trivial imprimitivity block that is invariant under the letter of defect 1.

5. Conclusion

We considered automata with an alphabet such that there is exactly one letter of defect 1 and the other letters are permutations over the state set. We found a necessary and sufficient condition to decide whether these automata are completely reachable. We saw that if the group generated by the permutations is primitive, then the automaton is completely reachable. On the other case, if the group is transitive and it has non trivial blocks of imprimitivity the condition depends on the behaviour of the letter of defect one over certain blocks of imprimitivity. The author believes that the additional condition stated in Theorem 3, the one stating the transitivity of the cores on the blocks of imprimitivity, can be omitted but more work on this direction must be done. In any case these results generalize what was presented in [7] where the alphabet was binary since the automata presented in that article are almost group and the group generated by the permutation letter is the cyclic one, which is abelian and thus the additional condition is given. Once decided whether or not an automaton is completely reachable, the next interesting question is to find a bound to the shortest word required to reach subsets of size $1 \leq k < n$. In [9] it is stated that this bound is at most $2n(n - k)$; but we believe that due to the strict structure of the considered automata the bound can be improved. Nevertheless this problem is open by the moment.

REFERENCES

1. Ananichev D., Vorel V. A new lower bound for reset threshold of binary synchronizing automata with sink. *J. Autom. Lang. Comb.*, 2019. Vol. 24, No. 2–4. P. 153–164. DOI: 10.25596/jalc-2019-153

2. André J. M. Near permutation semigroups. *Semigroups and Languages*, 2004. World Scientific. P. 43–53. DOI: 10.1142/9789812702616_0003

3. Berlinkov M. V., Nicaud C. Synchronizing almost-group automata. *Internat. J. Found. Comput. Sci.*, 2020. Vol. 31, No. 08. P. 1091–1112. DOI: 10.1142/S0129054120420058

4. Bondar E., Casas D., Volkov M. Completely reachable automata: An interplay between automata, graphs, and trees. *Internat. J. Found. Comput. Sci.*, 2023. Vol. 34, No. 06. P. 655–690. DOI: 10.1142/S0129054123450053

5. Bondar E. A., Volkov M. V. Completely reachable automata. In: *Lecture Notes in Comput. Sci., vol. 9777: Descriptional Complexity of Formal Systems (DCFS 2016), C. Câmpeanu, F. Manea, J. Shallit (eds)*. Cham: Springer, 2016. International Publishing. P. 1–17. DOI: 10.1007/978-3-319-41114-9_1

6. Bondar E. A., Volkov M. V. A characterization of completely reachable automata. In: *Lecture Notes in Comput. Sci., vol. 11088: Developments in Language Theory (DLT2018), M. Hoshi, S. Seki (eds)*. Cham: Springer, 2018. P. 145–155. DOI: 10.1007/978-3-319-98654-8_12

7. Casas D., Volkov M.V. Binary completely reachable automata. In: *Lecture Notes in Comput. Sci., vol. 13568: LATIN 2022: Theoretical Informatics, Castañeda A., Rodríguez-Henríquez F. (eds)*. Cham, 2022. Cham: Springer. P. 345–358. DOI: 10.1007/978-3-031-20624-5_21

8. Don H. The Černý conjecture and 1-contracting automata. *Electron. J. Combin.*, 2016. Vol. 23, No. 3. Art. no. 3.12. DOI: 10.37236/5616

9. Ferens R., Szykuła M. Completely reachable automata: a polynomial algorithm and quadratic upper bounds. In: *LIPIcs, vol. 261: 50th International Colloquium on Automata, Languages, and Programming (ICALP 2023), K. Etessami, U. Feige, G. Puppis (eds)*. Dagstuhl, Germany: Dagstuhl Publishing. P. 59:1–59:17. DOI: 10.4230/LIPIcs.ICALP.2023.59

10. Hoffmann S. New characterizations of primitive permutation groups with applications to synchronizing automata. *Inform. and Comput.*, 2023. Vol. 295(A). Art. no. 105086. DOI: 10.1016/j.ic.2023.105086

11. Maslennikova M.I. *Reset Complexity of Ideal Languages.* 2014. 12 p. arXiv:1404.2816v1 [cs.FL]

12. Rystsov I. Estimation of the length of reset words for automata with simple idempotents. *Cybernet. Systems Anal.*, 2000. Vol. 36, No. 3. P. 339–344. DOI: 10.1007/BF02732984

13. Rystsov I., Szykuła M. *Primitive Automata that are Synchronizing.* 2023. 13 p. arXiv:2307.01302v1 [cs.FL]