On Reversible Automata

Jean-Eric Pin

Bull Research and Development, Rue Jean-Jaurès, 78340 Les Clayes-sous-Bois, France

Abstract

A reversible automaton is a finite (possibly incomplete) automaton in which each letter induces a partial one-to-one map from the set of states into itself. We give four non-trivial characterizations of the languages accepted by a reversible automaton equipped with a set of initial and final states and we show that one can effectively decide whether a given rational (or regular) language can be accepted by a reversible automaton. The first characterization gives a description of the subsets of the free group accepted by a reversible automaton that is somewhat reminiscent of Kleene's theorem. The second characterization is more combinatorial in nature. The decidability follows from the third – algebraic – characterization. The last characterization relates reversible automata to the profinite group topology of the free monoid.

1 Introduction.

In this paper, we study a natural class of finite automata, the *reversible* automata, in which every letter induces a partial one-to-one map from the set of states into itself. More precisely, the aim of this paper is to describe effectively the languages accepted by these reversible automata. Although the statement of this problem requires only the very basic definitions of automata theory and could have been asked already in the fifties, the answer we propose is intimately related to the more advanced research on automata, finite semigroups and combinatorial group theory. A preliminary version of this paper has been presented in [22].

A (finite) automaton is a quintuple $\mathcal{A} = (Q, A, E, I, F)$ where Q is a (finite) set of states, A is a (finite) set of letters, called the *alphabet*, $E \subset Q \times A \times Q$ is the set of *edges*, $I \subset Q$ is the set of *initial* states and $F \subset Q$

is the set of *final states*. An edge (q, a, q') is also denoted $q \xrightarrow{a} q'$. A *path* in $\mathcal{A}$ is a finite sequence of consecutive edges :

$$p = q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \ \cdots \ q_{n-1} \xrightarrow{a_{n-1}} q_n$$

The *label* of the path p is the word $a_1 a_2 \ \cdots \ a_n$, its *origin* is q_0 and its *end* is q_n. A word is accepted by $\mathcal{A}$ if it is the label of a path in $\mathcal{A}$ having its origin in I and its end in F. The language (or set of words) accepted by $\mathcal{A}$ is denoted $|\mathcal{A}|$. An automaton is *deterministic* if it has a unique initial state and does not contain any pair of edges of the form $q \xrightarrow{a} q_1$ and $q \xrightarrow{a} q_2$ with $q_1 \neq q_2$.

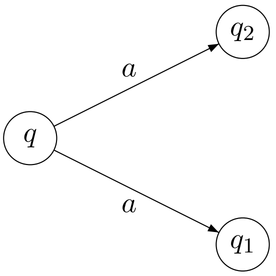


Figure 1.1: The forbidden configuration in a deterministic automaton.

In this case, every letter a induces a partial function from Q into itself, given by $q \to q \cdot a$, where $q \cdot a$ is the unique state q'(if it exists) such that $q \xrightarrow{a} q'$ is an edge in $\mathcal{A}$. If $u = a_1 a_2 \cdots a_n$ is a word, and q is a state, we set $q \cdot u = (\cdots((q \cdot a_1) \cdot a_2) \cdots) \cdot a_n$. This defines an action of A^* on Q and a monoid morphism $\varphi : A^* \to T(Q)$, where $T(Q)$ denotes the (finite) monoid of partial functions on Q under composition. Thus $\varphi(A^*)$ is a submonoid $M(\mathcal{A})$ of $T(E)$, called the *transition monoid* of $\mathcal{A}$. Every element $m \in M(\mathcal{A})$ is thus a partial function on Q. The action of m on a state q is also denoted $q \cdot m$.

Kleene's well-known theorem states that for a given language the three following conditions are equivalent : to be rational (or *regular*), to be accepted by a finite automaton or to be accepted by a finite deterministic automaton. Since rational languages are closed under reversal, the rational languages are also exactly the languages accepted by finite codeterministic automata. An automaton $\mathcal{A}$ is codeterministic if the reverse automaton $\mathcal{A}^r$ obtained by reversing the edges of $\mathcal{A}$ is deterministic. This is equivalent to saying that $\mathcal{A}$ contains a unique final state and does not contain any pair of edges of the form $q_1 \xrightarrow{a} q$ and $q_2 \xrightarrow{a} q$ with $q_1 \neq q_2$.

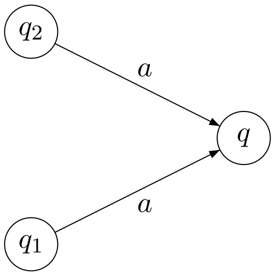


Figure 1.2: The forbidden configuration in a codeterministic automaton.

An automaton that contains neither the configuration given by Figure 1.1 nor the configuration given by figure 1.2 is said to be *reversible* (or *injective* [11, 27]). Thus an automaton is reversible if and only if each letter a induces a partial one-to-one map from the set of states into itself. The special case of reversible automata having a unique initial state and a unique final state has been considered in artificial intelligence in connection with the problem of inductively inferring general rules from examples [1]. They also have occurred in the study of the star-height problem [15] and are related to certain classes of biprefix codes [11]. The corresponding class of languages is not closed under union and the membership problem for this class is easy to solve.

Here we consider the general class of reversible automata, with no restriction on the sets of initial and final states. Now, the corresponding class of languages $\mathcal{C}$ is closed under (finite) union, but it is no longer trivial to decide whether or not a given rational language belongs to $\mathcal{C}$. For instance, the minimal automaton of a language of $\mathcal{C}$ is not reversible in general. That is, there exist languages that can be accepted by a reversible automaton but whose minimal automaton is not reversible. We propose in this paper four different characterizations of the class $\mathcal{C}$.

Our first characterization relates the class $\mathcal{C}$ to a class of subsets of the free group. Indeed, one can use reversible automata in a natural way to define subsets of the free group by considering an edge $q \xrightarrow{a} q'$ read "backwards" as an edge $q' \xrightarrow{\bar{a}} q$ with label $\bar{a}$, the formal inverse of the letter a in the free group. In this way, a reversible automaton accepts a subset $|\mathcal{A}|$ of A^* and a subset $||\mathcal{A}||$ of the free group such that $|\mathcal{A}| = ||\mathcal{A}|| \cap A^*$. Now the subsets of the free group accepted by a reversible automaton form the smallest class of subsets (of the free group) containing the singletons and closed under the three operations "union", "product by an element of the free group", and "subgroup generated by". These subsets are also the finite unions of cosets of finitely generated subgroups of the free group.

The other characterizations of $\mathcal{C}$ are based on a property of the syntactic monoid. Recall that the *syntactic monoid* of a subset L of A^* is the quotient $M(L)$ of A^* by the congruence $\sim_L$ defined by

$$u \sim_L v \text{ if and only if, for every } x, y \in A^*,\ xuy \in L \Leftrightarrow xvy \in L$$

It is also the transition monoid of the minimal automaton of L. The natural morphism $\eta : A^* \to M(L) = A^*/{\sim_L}$ is called the *syntactic morphism* and the subset $P = L\eta$ of $M(L)$ is called the *syntactic image* of L. It is well-known that a subset of A^* is rational if and only if its syntactic monoid is finite. Now, one can show that if L belongs to $\mathcal{C}$, then the idempotents of $M(L)$ commute. This property is not sufficient, however, to ensure that L belongs to $\mathcal{C}$. There are three different ways to strengthen this condition to obtain a characterization of $\mathcal{C}$. The first solution is to require the additional condition that in L, “plus is equivalent to star”, or more precisely, that, if $xu^+y \subset L$ for some words $x, u, y \in A^*$, then $xu^*y \subset L$. This gives our first characterization. However, this characterization is not fully satisfactory because it is not clear whether there is an algorithm to verify this additional condition. We shall treat this problem in detail in section 6 and give a polynomial algorithm for testing whether a language given by a finite deterministic automaton belongs to $\mathcal{C}$. In particular the membership problem for $\mathcal{C}$ is decidable for rational languages. In fact, one can also give a purely algebraic (and effective) characterization of $\mathcal{C}$: a language L belongs to $\mathcal{C}$ if and only if the idempotents of $M(L)$ commute and the syntactic image P of L satisfies the following condition : for every $s, t \in M(L)$, and for every idempotent e of $M(L)$,

$$set \in P \text{ implies } st \in P. \qquad (*)$$

Our last characterization relates $\mathcal{C}$ to the profinite group topology of the free monoid [10, 20, 27]. This topology is defined by a distance in which, roughly speaking, two words are close if they are not distinguishable by a group of small cardinality. We show that a rational language L belongs to $\mathcal{C}$ if and only if the idempotent of $M(L)$ commute and L is *closed* in this topology. In fact, it has recently been proved, as the conclusion of a cascade of partial results, that a rational language L is closed if and only if its syntactic image satisfies Condition $(*)$. This is a very nice example of an algebraic characterization of a topological property. It also gives a simple algorithm for computing the closure of a given rational language. See [25, 26, 12] for more details.

The paper is organized as follows. Section 2 contains some basic facts about reversible automata. The connections with the free group are presented in section 3. The algebraic characterizations are given in section 4 and the topological aspects of the problem are discussed in section 5. Section 6 is devoted to algorithms and section 7 is a concluding section.

2 Reversible versus bideterministic automata.

A reversible automaton equipped with a unique initial state and a unique final state is called *bideterministic*. Thus an automaton is bideterministic if and only if it is both deterministic and codeterministic.

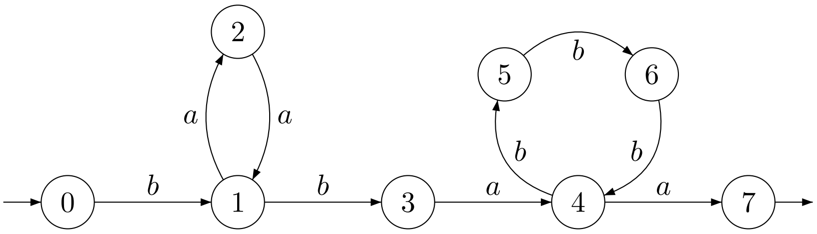


Figure 2.3: A bideterministic automaton.

It is not difficult to see that a trim bideterministic automaton is necessarily minimal. The next proposition, discovered independently by various authors, characterizes the languages accepted by bideterministic automata.

Proposition 2.1 *A language L is accepted by a bideterministic automaton if and only if the minimal automaton of L is reversible and has a unique final state.*

It is much more difficult to characterize the class $\mathcal{C}$ of all the languages accepted by a reversible automaton.

Example 2.1 The automaton represented in the following diagram is reversible and accepts the language $\{a, ac, bc\}$. This automaton is not connected.

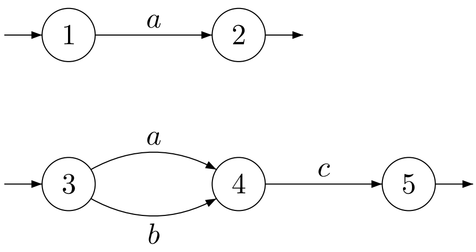


Figure 2.4: A reversible automaton accepting $\{a, ac, bc\}$.

The following construction shows that $\mathcal{C}$ is closed under union. Given two reversible automata $\mathcal{A}_1 = (Q_1, A, E_1, I_1, F_1)$ and $\mathcal{A}_2 = (Q_2, A, E_2, I_2, F_2)$, we form the *disjoint union* $\mathcal{A}$ of $\mathcal{A}_1$ and $\mathcal{A}_2$ as follows : $\mathcal{A} = (Q, A, E, I, F)$ where Q (resp. E, I, F) is the disjoint union of Q_1 and Q_2 (resp. E_1 and E_2, I_1 and I_2, F_1 and F_2). Then $\mathcal{A}$ is a reversible automaton that accepts $L(\mathcal{A}_1) \cup L(\mathcal{A}_2)$. It follows that a language is accepted by a reversible automaton if and only if it is a finite union of languages accepted by bideterministic automata.

Since the minimal automaton of a singleton $\{u\}$ is always reversible, it follows that $\mathcal{C}$ contains all the finite languages. Notice that the minimal automaton of a language of $\mathcal{C}$ is not necessarily reversible. For instance, the minimal automaton of $L = \{a, ac, bc\}$ is not reversible.

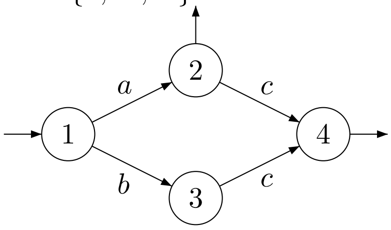


Figure 2.5: The minimal automaton of $\{a, ac, bc\}$.

3 Reversible automata in the free group.

First recall the definition of the free group on a set A. Let $\tilde{A} = A \cup \{\bar{a} \mid a \in A\}$. The free group $FG(A)$ is the quotient of $\tilde{A}^*$ by the congruence generated by the relations $a\bar{a} = \bar{a}a = 1$ for every $a \in A$. We denote $\pi : \tilde{A}^* \to FG(A)$ the natural morphism. We need first to define the subset of the free group accepted by an automaton. Let $\mathcal{A} = (Q, A, E, I, F)$ be an automaton. We form the automaton $\tilde{\mathcal{A}} = (Q, \tilde{A}, \tilde{E}, I, F)$ by setting

$$\tilde{E} = E \cup \{(q', \bar{a}, q) \mid (q, a, q') \in E\}$$

Thus, intuitively, to each edge $q \xrightarrow{a} q'$ is attached a "reverse" edge $q' \xrightarrow{\bar{a}} q$ whose label is the formal inverse of a. Now the label of a path in $\tilde{\mathcal{A}}$ is a word u in $(\tilde{A})^*$ that defines the element $u\pi$ of the free group. By definition, the subset of the free group accepted by $\mathcal{A}$ is the set

$$||\mathcal{A}|| = |\tilde{\mathcal{A}}|\pi$$

For instance, if $\mathcal{A}$ is the automaton of example 2.1, then $||\mathcal{A}|| = \{a\} \cup \langle a\bar{b}\rangle\{ac, bc\}$ where $\langle X\rangle$ denotes the subgroup of $FG(A)$ generated by a set X.

Since $|\mathcal{A}| = ||\mathcal{A}|| \cap A^*$, it suffices to describe the subsets of the free group accepted by a reversible automaton to obtain a first characterization of the class $\mathcal{C}$. We first recall a definition. The *rational subsets* of the free group $FG(A)$ form the smallest class $\mathcal{R}$ of subsets of $FG(A)$ such that:

(a) every finite subset of $FG(A)$ belongs to $\mathcal{R}$,

(b) if S and T are in $\mathcal{R}$, then so are ST and $S \cup T$,

(c) if S is in $\mathcal{R}$, then so is S^*, the submonoid of $\mathcal{R}$ generated by S.

It is easy to show that if S is rational, then $\langle S\rangle$ is also rational. In fact, the rational subgroups of the free group can be characterized as follows (see [5] for a proof).

Proposition 3.1 *A subgroup of $FG(A)$ is rational if and only if it is finitely generated.*

We can now state our first characterization.

Theorem 3.2 *A subset S of the free group $FG(A)$ is accepted by a reversible automaton if and only if S is a finite union of left cosets of finitely generated subgroups of the free group.*

Proof. Let $\mathcal{A} = (Q, A, E, I, F)$ be a reversible automaton. For every $p, q \in Q$, set $\mathcal{A}_{p,q} = (Q, A, E, \{p\}, \{q\})$. Then $||\mathcal{A}|| = \cup_{p\in I, q\in F}||\mathcal{A}_{p,q}||$ and, if g is any element of $||\mathcal{A}_{p,q}||$, it is easy to see that $||\mathcal{A}_{p,q}|| = g||\mathcal{A}_{q,q}||$. Furthermore, since $||\mathcal{A}_{q,q}|| = |\tilde{\mathcal{A}}_{q,q}|\pi$, each $||\mathcal{A}_{q,q}||$ is the image of a rational subset of $\tilde{A}^*$ and thus is a rational group. By proposition 3.1, it follows that $||\mathcal{A}_{q,q}||$ is finitely generated, and thus $||\mathcal{A}||$ is a finite union of left cosets of finitely generated subgroups of the free group.

Conversely, the class of subsets of the free group accepted by a reversible automaton is closed under finite union: if S_i is accepted by $\mathcal{A}_i$ for $1 \leq i \leq$

n, then the disjoint union of the $\mathcal{A}_i$'s accepts the set $\cup_{1\leq i\leq n} S_i$. Next, if (Q, A, E, I, F) is a reversible automaton, and if $g \in FG(A)$, then

$$g||(Q, A, E, I, F)|| = ||(Q, A, E, I', F)||$$

where $I' = I \cdot g^{-1}$ is the set of states such that there exists a path with label g from q to a state of I. Finally, it suffices to show that every finitely generated subgroups of the free group is accepted by a reversible automaton. Let H be a finitely generated subgroup of $FG(A)$ and let $\{h_1, \ldots, h_n\}$ be a set of generators of H. Fix some words $u_i \in \tilde{A}^*$ such that $u_i\pi = g_i$ and set $U = \{u_1, \ldots u_n\}$. Consider the "flower" automaton associated with the u_i's. This is the automaton $\mathcal{B} = (Q, A, E, \{1\}, \{1\})$, where $Q = \{1\} \cup \{(x, y) \in \tilde{A}^+ \mid xy \in U\}$ and $E = E_1 \cup E_2 \cup E_3$, where

$$\begin{aligned}
E_1 &= \{(1, a, (a, y)) \mid a \in A,\ ay \in U\} \cup \{((\bar{a}, y), a, 1) \mid a \in A,\ \bar{a}y \in U\} \\
E_2 &= \{((x, a), a, 1) \mid a \in A,\ xa \in U\} \cup \{(1, a, (x, \bar{a})) \mid a \in A,\ x\bar{a} \in U\} \\
E_3 &= \{((x, ay), a, (xa, y)) \mid a \in A,\ xay \in U\} \\
&\quad \cup \{((xa, y), a, (x, ay)) \mid a \in A,\ x\bar{a}y \in U\}
\end{aligned}$$

Now the automaton $\tilde{\mathcal{B}}$ (obtained from $\mathcal{B}$ by adding the reverse edges, as explained above) accepts the language $\{u_1, u_2, \ldots, u_n, \bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n\}^*$ and thus $||\mathcal{B}|| = |\tilde{\mathcal{B}}|\pi = H$. This automaton is not reversible in general but can be converted into a reversible automaton accepting H by the following process. Each step of the process consists in identifying the states q_1 and q_2 appearing in one of the configurations given by figures 1.1 and 1.2. This transforms $\mathcal{B}$ into an automaton $\mathcal{B}'$ such that $|\mathcal{B}| \neq |\mathcal{B}'|$ in general, but such that $||\mathcal{B}|| = ||\mathcal{B}'||$. The reason is that if $q_1 \xrightarrow{a} q$ and $q_2 \xrightarrow{a} q$ (or symmetrically, $q \xrightarrow{a} q_1$ and $q \xrightarrow{a} q_2$) are two edges of $\mathcal{B}$, then there is a path of label $a\bar{a}$ (resp. $\bar{a}a$) between q_1 and q_2 in $\tilde{\mathcal{B}}$. Since the number of states decreases at each step, one obtains a reversible automaton after a finite number of steps. One can show that the result does not depend on the order in which the identifications are made. □

The previous construction is illustrated in the following diagram, where $H = \langle a\bar{b}a, abba \rangle$. At each step, the hatchured states are identified.

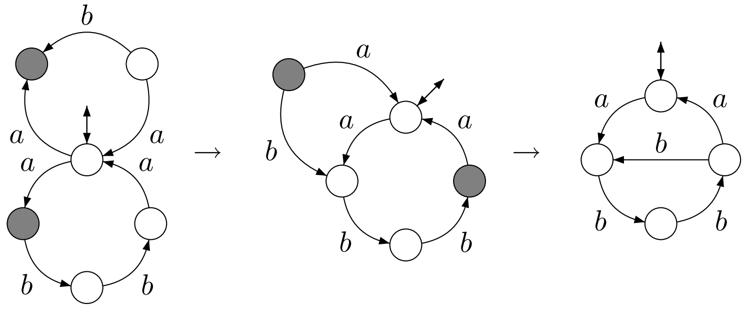


Figure 3.6: Construction of a reversible automaton accepting $\langle a\bar{b}a, abba\rangle$.

Here is another version of Theorem 3.2, that is somewhat reminiscent of Kleene's theorem.

Theorem 3.3 *The subsets of the free group accepted by a reversible automaton form the smallest class of subsets* $\mathcal{F}$ *such that*

(1) $\emptyset \in \mathcal{F}$ *and for every* $g \in FG(A)$, $\{g\} \in \mathcal{F}$,

(2) *if* $S_1, S_2 \in \mathcal{F}$, *then* $S_1 \cup S_2 \in \mathcal{F}$,

(3) *if* $S \in \mathcal{F}$ *and* $g \in FG(A)$ *then* $gS \in \mathcal{F}$,

(4) *if* $S \in \mathcal{F}$, *then* $\langle S\rangle \in \mathcal{F}$.

Proof. Let $\mathcal{S}$ be the class of all subsets of the free group accepted by a reversible automaton. By theorem 3.2, $\mathcal{S}$ is also the class of all subsets of $FG(A)$ that are finite unions of left cosets of finitely generated subgroups of the free group. By proposition 3.1, every finitely generated subgroup is rational and thus every element of $\mathcal{S}$ is rational. Now, $\mathcal{S}$ contains the singletons and is closed under finite union and under the operation $S \to gS$ for every element $g \in FG(A)$. Finally, if $S \in \mathcal{S}$, then S is rational and so is $\langle S\rangle$. By proposition 3.1, $\langle S\rangle$ is finitely generated and thus belongs to $\mathcal{S}$. It follows that $\mathcal{S}$ satisfies properties (1)-(4) and thus $\mathcal{F} \subset \mathcal{S}$. Conversely, since $\mathcal{F}$ is closed under finite union and contains the singletons, it contains the finite sets and hence, by (4), the finitely generated subgroups of $FG(A)$. Finally it contains the cosets of finitely generated subgroups by (3) and the class $\mathcal{S}$ by (2). Thus $\mathcal{F} = \mathcal{S}$. □

4 An algebraic characterization of $\mathcal{C}$.

Let L be a rational language of A^*. We denote by $M(L)$ the syntactic monoid of L, by $\eta : A^* \to M(L)$ the syntactic morphism and by $P = L\eta$ the syntactic image of L. Recall that an element e of a monoid M is *idempotent* if $e = e^2$. The following proposition gives two important properties of the languages of $\mathcal{C}$.

Proposition 4.1 *If L is accepted by a reversible automaton, then,*

(a) *the idempotents of $M(L)$ commute,*

(b) *for every $x, u, y \in A^*$, $xu^+y \subset L$ implies $xy \in L$.*

Proof. Let $\mathcal{A} = (Q, A, E, I, F)$ be a reversible automaton accepting L. Let M be the transition monoid of $\mathcal{A}$. Let e be an idempotent of M. Then for every $q \in Q$, $(q \cdot e) \cdot e = q \cdot e$ whenever $q \cdot e$ is defined. Since $\mathcal{A}$ is reversible, e is an injective partial function and thus $q \cdot e = q$ or is undefined. In other words, every idempotent is a subidentity on Q. It follows immediately that the idempotents commute in M. Now, the syntactic monoid $M(L)$ divides M (see [9, 13, 21] for instance), and since the class of monoids with commuting idempotents is closed under division, the idempotents also commute in $M(L)$.

Let $x, u, y \in A^*$ be words such that $xu^+y \subset L$. Since M is finite, there exists an integer $n > 0$ such that u^n is idempotent in M, that is, induces a subidentity on Q. Now since $xu^+y \subset L$, there exists an initial state q and a final state q' such that $q \cdot xu^ny = q'$. Thus $(q \cdot x) \cdot u^n$ is defined, and hence is equal to $(q \cdot x)$. Therefore $q \cdot xy = q \cdot xu^ny = q'$ whence xy is accepted by $\mathcal{A}$ and thus $xy \in L$. □

Condition (b) of the previous proposition is equivalent to a more algebraic statement.

Proposition 4.2 *For every rational language L, the following conditions are equivalent:*

(b) *for every $x, u, y \in A^*$, $xu^+y \subset L$ implies $xy \in L$,*

(c) *for every $s, t \in M(L)$, and for every idempotent $e \in M(L)$, $set \in P$ implies $st \in P$.*

Proof. Assume that (b) is satisfied and let $s, e, t \in M(L)$ with e idempotent. Assume that $set \in P$. Then, since η is surjective, there exist some words $x, u, y \in A^*$ such that $x\eta = s$, $u\eta = e$ and $y\eta = t$. Now, for every $n > 0$,

$(xu^ny)\eta = set \in P$. Thus $xu^+y \in P\eta^{-1} = L$, and hence $xy \in L$ by (b). It follows that $st = (xy)\eta \in L\eta = P$.

Conversely, assume that (c) holds, and let x, u, y be words such that $xu^+y \subset L$. Then there exists $n > 0$ such that $u^n = e$ is an idempotent. Setting $x\eta = s$ and $y\eta = t$, we obtain $set \in L\eta = P$, and hence $st \in P$ by (c). Therefore $xy \in L$, since $(xy)\eta = st \in P$. □

We now turn to the converse of proposition 4.1, for which we need a more detailed study of monoids with commuting idempotents. Recall that an element x of a monoid M is *regular* if there exists an element y such that $xyx = x$ and $yxy = y$. We start with an important combinatorial lemma, due to Ash [2].

Proposition 4.3 *Let M be a monoid with commuting idempotents, and let $\eta : A^* \to M$ be a monoid morphism. Then there exists an integer $N > 0$ such that every word $w \in A^*$ admits a factorization of the form $w = u_0v_1u_1...v_ku_k$ with $u_1, \ldots, u_{k-1} \in A^+$, $u_0, u_k \in A^*$, $v_1, \ldots, v_k \in A^+$ and*

(1) *$v_1\eta$, ..., $v_k\eta$ are regular elements of M,*

(2) *if b_{i-1} denotes the last letter of u_{i-1} and a_i the first letter of u_i, $(b_iv_i)\eta$ and $(v_ia_i)\eta$ are not regular,*

(3) *$|u_0...u_k| \leq N$.*

Recall the definition of Green's relations $\mathcal{R}$ and $\mathcal{L}$. Let u and v be two elements of a monoid M. Then $u \mathrel{\mathcal{R}} v$ (resp. $u \mathrel{\mathcal{L}} v$) if and only if there exist two elements $x, y \in M$ such that $ux = v$ and $vy = x$ (resp. $xu = v$ and $yv = x$). An $\mathcal{R}$-class (resp. $\mathcal{L}$-class) is regular if it contains a regular element. One can show [21] that an element m of a monoid M is regular if and only if its $\mathcal{R}$-class (resp. $\mathcal{L}$-class) is regular, or equivalently, contains an idempotent. The next proposition summarizes the properties of $\mathcal{R}$-classes that are used in this paper.

Proposition 4.4 *Let M be a monoid with commuting idempotents. Then*

(1) *every regular $\mathcal{R}$-class R contains a unique idempotent e,*

(2) *for every $x \in R$, $ex = x$,*

(3) *for every $u, v, s \in M$, $u \mathrel{\mathcal{R}} v \mathrel{\mathcal{R}} us$ and $us = vs$ implies $u = v$.*

Proof. Since R is regular, it contains an idempotent e. Let $x \in R$. Then $e \mathrel{\mathcal{R}} x$ and thus $e = xy$ and $x = ez$ for some $y, z \in M$. Now $ex = eez = ez = x$, which proves (2). In particular, if f is another idempotent of R, one has $ef = f$ and $fe = e$. Since the idempotents commute by assumption, it follows $ef = fe = e = f$, proving (1).

We need a little bit more of semigroup theory to prove (3). Since $u \mathrel{\mathcal{R}} us$ and $u \mathrel{\mathcal{R}} v$, there exist $t, a \in M$ such that $u = ust$ and $v = ua$. Let G be the minimal ideal of the semigroup $S = \{x \in M \mid ux = u\}$. It is a well-known fact of semigroup theory that the minimal ideal of a semigroup in which the idempotents commute is actually a group whose identity is an idempotent f. Therefore $uf = u = ust = ustf$ and $ufastf = uastf = vstf = ustf = u$. It follows $stf, fastf \in G$ and since G is a group, $fa \in G$, that is $ufa = u$. But $ufa = ua = v$ and thus $u = v$. □

Theorem 4.5 *A rational language L is accepted by a reversible automaton if and only if it satisfies conditions (a) and (b) or, equivalently, conditions (a) and (c).*

Proof. By propositions 4.1 and 4.2, it suffices to show that if L satisfies (a) and (c), then $L \in \mathcal{C}$. Let r be the maximum size of an $\mathcal{R}$-class of $M(L)$, and let N be the integer given by proposition 4.3. Let $\mathcal{F}$ be the set of all the reversible automata of the form $\mathcal{B} = (Q, A, E, I, F)$ where Q contains at most $r(N+1)$ states and the language accepted by $\mathcal{B}$ is contained in L. $\mathcal{F}$ is a finite set, since there are only a finite number of automata with at most $r(N+1)$ states. Let $\mathcal{A}$ be the disjoint union of all the automata of $\mathcal{F}$. Then $\mathcal{A}$ is a reversible automaton such that $L(\mathcal{A}) \subset L$. To prove that $L(\mathcal{A})$ is actually equal to L, it suffices to exhibit, for every word $w \in L$, a reversible automaton $\mathcal{B}$ of $\mathcal{F}$ such that $w \in L(\mathcal{B})$.

Let $m = w\eta$ and denote by $P(m)$ the smallest subset of $M(L)$ containing m and satisfying condition (c): for every $s, e, t \in M(L)$, with e idempotent, $set \in P(m)$ implies $st \in P(m)$. Now, since $m \in P$, $P(m)$ is contained in P and the language $L(m) = P(m)\eta^{-1}$ is contained in L.

We first assume that m is a regular element of $M(L)$. Then, by proposition 4.4, the $\mathcal{R}$-class R of m contains a unique idempotent e and we can state

Lemma 4.6 *The language $L(m)$ is accepted by the reversible automaton $\mathcal{B} = (R, A, E, \{e\}, \{m\})$, where $E = \{(x, a, x{\cdot}a\eta) \mid x \in R, a \in A$ and $x(a\eta) \in R\}$.*

Proof. Proposition 4.4 shows that $\mathcal{B}$ is reversible. Next, $L(\mathcal{B}) = S\eta^{-1}$ where $S = \{\ s \in M(L) \mid es = m\ \}$ and $L(m) = P(m)\eta^{-1}$. Therefore, it suffices to show that $S = P(m)$. First $m \in S$ by proposition 4.4, and if

$sft \in S$ for some $s, t \in M(L)$ and some idempotent f, then $es \ \mathcal{R} \ esf \ \mathcal{R} \ (es)f = (esf)f$. It follows that $es = esf$ by proposition 4.4, whence $est = esft = m$ and $st \in S$. Thus S satisfies condition (c) and $P(m)$ is contained in S. Conversely, let $s \in S$. Then $es = m$, and hence $1es = m \in P(m)$. Therefore, by condition (c), $s = 1.s \in P(m)$ and thus $S = P(m)$ as required.

We now turn to the general case. Let $w = u_0 v_1 u_1 \cdots v_k u_k$ be a factorization of w given by proposition 4.3. For $1 \leq i \leq k$, let $v_i\eta = m_i$, and let e_i be the (unique) idempotent of the $\mathcal{R}$-class of m_i. The previous proposition shows that the language $L(m_i)$ is accepted by the automaton $\mathcal{B}_i = (R_i, A, \cdot, e_i, m_i)$.

We consider also the minimal automaton $\mathcal{B}$ of the word $u = u_0 u_1 \cdots u_k$ defined as follows. The set of states is the set of left factors of u and, for each letter $a \in A$ and for each left factor x of u, $x \cdot a = xa$ if xa is a left factor of u and is undefined otherwise. We now “sew” the automata $\mathcal{B}$ and $\mathcal{B}_i$'s together, according to the following diagram.

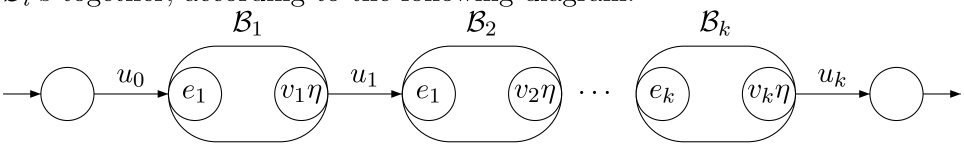


Figure 4.7: Sewing $\mathcal{B}$ and the $\mathcal{B}_i$'s together

Now, proposition 4.3 implies that the resulting automaton is reversible (the details are omitted), accepts the language

$$K = u_0 L(m_1) u_1 \cdots u_{k-1} L(m_k) u_k$$

and contains at most $r(N+1)$ states. We claim that K is contained in $L(m)$ (and thus in L). Indeed put, for $0 \leq i \leq k$, $s_i = u_i\eta$, so that $m = s_0 m_1 s_1 \cdots m_k s_k$. Since $K \subset K\eta\eta^{-1}$, it suffices to show that $K\eta = s_0 P(m_1) s_1 \cdots P(m_k) s_k$ is contained in $P(m)$. Let T be the set of all $(t_1, \ldots, t_k)$ of $P(m_1) \times \cdots \times P(m_k)$ such that $s_0 t_1 s_1 \cdots s_k t_k \in P(m)$. Then T contains $(m_1, \cdots, m_k)$. Furthermore, if $(t_1, \cdots, t_k) \in T$ and if $t_i = x_i f_i y_i$ for some idempotent f_i, then $(s_0 t_1 \cdots s_{i-1} x_i) f_i (y_i s_i ... s_k t_k) \in P(m)$, and hence, by condition (c), $s_0 t_1 \cdots s_{i-1} x_i y_i s_i ... s_k t_k \in P(m)$, so that $(t_1, \ldots, t_{i-1}, x_i y_i, t_{i+1}, \ldots, t_k) \in T$. Therefore T is equal to $P(m_1) \times \cdots \times P(m_k)$ and this concludes the proof. □

5 A topological characterization.

In this section, we give a topological description of the class $\mathcal{C}$. Let us first define the profinite group topology. One can show that two distinct words u and v of A^* can always be separated by a finite group in the following sense: there exists a finite group G and a monoid morphism $\varphi : A^* \to G$ such that $\varphi(u) \neq \varphi(v)$. Set, for every $u, v \in FG(A)$,

$$r(u,v) = \min \{ \ Card(G) \mid G \text{ is a finite group that separates } u \text{ and } v \ \}$$

and

$$d(u,v) = e^{-r(u,v)}$$

with the usual conventions $\min \emptyset = \infty$ and $e^{-\infty} = 0$. Then d is a distance (in fact an ultrametric distance) which defines a topology on A^*, called the *profinite group topology* of the free monoid. This topology, introduced by Reutenauer [27, 28], is an analogue for the free monoid to the profinite topology of the free group introduced by M. Hall [10]. It is the coarsest topology such that every monoid morphism from A^* into a discrete finite group is continuous. The free monoid A^*, equipped with this topology, is a topological monoid. The interested reader is referred to [20, 27] for a more detailed study of this topology. An example of a converging sequence is given by the following proposition, proved in [27].

Proposition 5.1 *For every word $w \in A^*$, $\lim\limits_{n \to \infty} w^{n!} = 1$.*

The next proposition relates reversible automata to this topology.

Proposition 5.2 [27] *Every language accepted by a reversible automaton is closed in the profinite group topology.*

The converse is not true in general. For instance, the language a^*b^* is closed but is not accepted by any reversible automaton. However, we have

Theorem 5.3 *A rational language L is accepted by a reversible automaton if and only if the idempotents commute in $M(L)$ and L is closed in the profinite group topology.*

Proof. If L is accepted by a reversible automaton, then L is closed by proposition 5.2 and the idempotents of $M(L)$ commute by proposition 4.1. Conversely, if L is closed, then L satisfies (b). Indeed, let x, u, y be words such that $xu^+y \subset L$. Then, in particular, for every $n > 0$, $xu^{n!}y \in L$.

Since L is closed, and since the multiplication is continuous, it follows by Proposition 5.1 that $xy = \lim_{n\to\infty} xu^{n!}y \in L$. Thus L satisfies (a) and (b) and the result follows from theorem 4.5. □

In fact, one can characterize the closed languages in the same way.

Theorem 5.4 *A rational language of A^* is closed if and only if it satisfies condition (c) (or (b)).*

The history of this result is quite interesting. It was first conjectured by the author in [20, 25], and was shown to be equivalent with the Rhodes "Type II" conjecture in semigroup theory (the author [25, 23] in one direction, and Margolis and the author [19] in the other direction). It was also shown to be a consequence of another conjecture on the profinite topology of the free group, proposed by Reutenauer and the author [26] :

Let $H_1, \ldots, H_n$ be finitely generated subgroups of $FG(A)$. Then $H_1H_2\cdots H_n$ is closed in the profinite topology.

Then Ash proved the Rhodes conjecture [3, 4], giving as a byproduct the first (rather indirect) proof of theorem 5.4. Next, Margolis [16] showed that the two topological conjectures where actually equivalent, giving thus the first proof of the topological conjecture for the free group. Finally, Ribes and Zalesskii [29] gave a direct proof of the topological conjecture for the free group, giving in turn another proof of theorem 5.4.

6 Algorithms

In this section, we give a polynomial time algorithm for testing, given an n-state deterministic automaton $\mathcal{A}$, whether $|\mathcal{A}|$ belongs to $\mathcal{C}$ or not. First we may assume that $\mathcal{A}$ is a complete, minimal, deterministic automaton, since completion and minimalization can be achieved in polynomial time and do not increase the number of states by more than one.

Before giving the details of our algorithm, let us fix some convenient notation. Given a finite (complete) deterministic automaton $\mathcal{A} = (Q, A, \cdot)$ and a positive integer k, we denote by $\mathcal{A}^k = (Q^k, A, \cdot)$ the direct product of k copies of $\mathcal{A}$, where the action of A on Q^k is given by $(q_1, \ldots, q_k) \cdot a = (q_1 \cdot a, \ldots, q_k \cdot a)$. We also denote by $G_k(\mathcal{A})$ the transitive closure of the directed graph defined by $\mathcal{A}^k$. This construction is illustrated in the figure below.

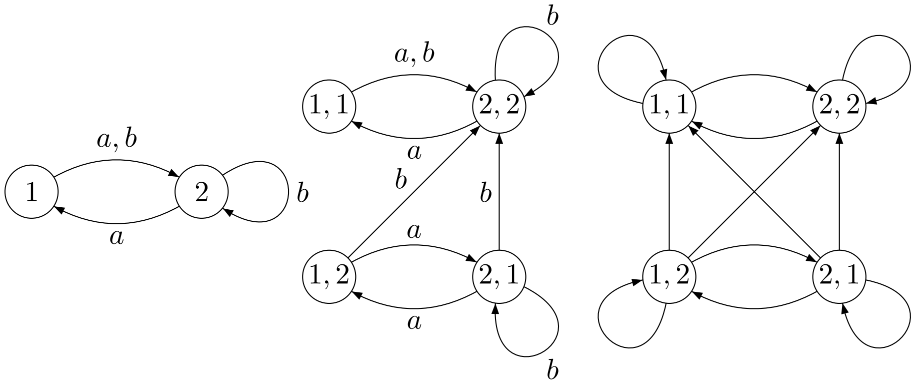
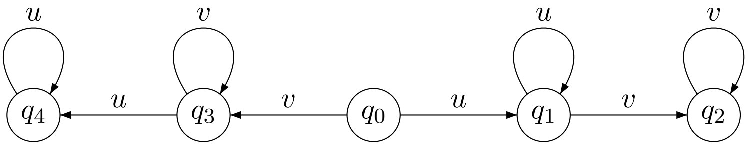


Figure 6.8: An automata $\mathcal{A}$, the automaton $\mathcal{A}^2$ and the graph $G_2(\mathcal{A})$.

Given a deterministic automaton $\mathcal{A} = (Q, A, \cdot)$, the set of all paths in $\mathcal{A}$ defines an infinite labelled graph $G(\mathcal{A})$, with Q as set of vertices, and the triples of the form $(q, w, q.w)$ (where $w \in A^+$) as edges. A labelled subgraph of $G(\mathcal{A})$ is said to be a *configuration* present in $\mathcal{A}$.

We first give a polynomial algorithm for testing whether the idempotents of $M(|\mathcal{A}|)$ commute.

Theorem 6.1 *Let $\mathcal{A}$ be the minimal automaton of a language L. The idempotents of the syntactic monoid of L commute if and only if there exist no configuration of $\mathcal{A}$ of the form*



with $q_2 \neq q_4$.

Proof. Let $M = M(L)$ and let $\eta : A^* \to M$ be the natural morphism. First assume that the idempotents of M commute. If $\mathcal{A}$ contains the above configuration, we have for every $n > 0$,

$$q_0 \cdot u^n v^n = q_2 \quad \text{and} \quad q_0 \cdot v^n u^n = q_4$$

In particular, since M is a finite monoid, one can choose n such that $u^n\eta$ and $v^n\eta$ are idempotent. Since idempotents commute in M, we obtain $q_2 = q_4$.

Conversely, assume that $\mathcal{A}$ does not contain the above configuration, and let e, f be two idempotents of M. Choose $u, v \in A^*$ such that $u\eta = e$ and $v\eta = f$. We claim that for any state $q_0 \in Q$, $q_0 \cdot uv = q_0 \cdot vu$. Indeed, set

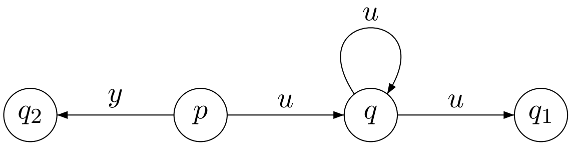
$q_1 = q_0 \cdot u$, $q_2 = q_1 \cdot v$, $q_3 = q_0 \cdot v$ and $q_4 = q_3 \cdot u$. Then since $u\eta$ and $v\eta$ are idempotent, one has $q_1 \cdot u = q_0 \cdot u^2 = q_0 \cdot u = q_1$ and, similarly, $q_2 \cdot v = q_2$, $q_3 \cdot v = q_3$ and $q_4 \cdot u = q_4$. Thus we have a forbidden configuration unless $q_2 = q_4$, that is $q_0 \cdot uv = q_0 \cdot vu$. □

Corollary 6.2 *There is a polynomial time algorithm for testing whether the idempotents of the syntactic monoid of $|\mathcal{A}|$ commute.*

Proof. It suffices to compute $G_4(\mathcal{A})$ and to check whether it contains edges of the form $\big((q_0, q_1, q_3, q_4), (q_1, q_1, q_4, q_4)\big)$ and $\big((q_0, q_1, q_2, q_3), (q_3, q_2, q_2, q_3)\big)$ with $q_2 \neq q_4$. □

We now turn to the second condition which characterizes the class $\mathcal{C}$, the condition (*).

Theorem 6.3 *Let $\mathcal{A} = (Q, A, E, \{i\}, F)$ be the minimal automaton of a language L. Then L satisfies the condition (*) if and only if there exist no configuration of $\mathcal{A}$ of the form*



with $q_1 \in F$ and $q_2 \notin F$.

Proof. Suppose that L satisfies (*), and consider a configuration in $\mathcal{A}$ of the form above. Since $\mathcal{A}$ is minimal, every state of $\mathcal{A}$ is accessible and in particular, there exists a word $x \in A^*$ such that $i \cdot x = p$. It follows, for every $n > 0$, $i \cdot xu^n y = p \cdot u^n y = q \cdot y = q_1 \in F$. Therefore $xy \in L$, that is $i \cdot xy = p \cdot y = q_2 \in F$.

Conversely, suppose that $\mathcal{A}$ has no configuration of the form above. Assume that for some $x, y, u \in A^*$, $xu^+y \subset L$. Let n be an integer such that $u^n\eta$ is idempotent. Set $v = u^n$, $p = i \cdot x$, $q = p \cdot v$ and $q_1 = q \cdot y$. Then $q \cdot v = p \cdot v^2 = p \cdot v = q$. Furthermore $q_1 = i \cdot xvy \in F$ since $xvy \in L$. Thus $p \cdot y \in F$, otherwise $\mathcal{A}$ would contain a forbidden configuration. □

Corollary 6.4 *There is a polynomial time algorithm for testing whether the language accepted by an n-state minimal automaton is closed.*

Corollary 6.5 *There is a polynomial time algorithm for testing whether the language accepted by an n-state minimal automaton can be accepted by a reversible automaton.*

7 Conclusion.

Let us summarize the four characterizations of the languages accepted by a reversible automaton into a single statement.

Theorem 7.1 *Let L be a rational language. Let M (resp. P) be its syntactic monoid (resp. image). The following conditions are equivalent:*

(1) *L is accepted by a reversible automaton,*

(2) *$L = K \cap A^*$ where K is a subset of the free group $FG(A)$ consisting of a finite union of left cosets of finitely generated subgroups of $FG(A)$,*

(3) *the idempotents of M commute and, for every $x, u, y \in A^*$, $xu^+y \in L$ implies $xy \in L$,*

(4) *the idempotents of M commute and, for every $s, t, e \in M$ such that e is idempotent, $set \in P$ implies $st \in P$,*

(5) *the idempotents of M commute and L is closed in the profinite group topology of A^*.*

We have also given a polynomial time algorithm for testing whether the language accepted by a given n-state deterministic automaton can be accepted by a reversible automaton. This algorithm does not give, however, any bound on the number of states of the smallest reversible automaton accepting the language. More precisely, given a language $L \in \mathcal{C}$, denote by $m(L)$ the number of states of its minimal automaton and by $c(L)$ the number of states of a smallest reversible automaton accepting L. It may happen that $c(L) < m(L)$, because one can have several initial states in a reversible automaton. It would be interesting to estimate the functions

$$r(n) = \min\{c(L) \mid m(L) = n\} \quad \text{and} \quad R(n) = \max\{c(L) \mid m(L) = n\}$$

Here is a first estimation.

Proposition 7.2 $r(n) = O\big(\frac{\ln n}{\ln \ln n}\big)^2$

Proof. We exhibit, for every $n > 2$, a language L_n such that $c(L_n) \leq n^2$ and $m(L_n) = n!$. This leads to the result by a simple application of Stirling's formula. Our construction is adapted from a construction of Birget [6].

Let $A = \{a, b\}$ and $Q_n = \{0, 1, \ldots, n-1\}$. Let a act on Q_n as the cyclic permutation $(0, 1, \ldots, n-1)$ and b as the transposition $(0, 1)$. Finally, let E be the set of edges defined by this action : $E = \{(q, a, q \cdot a) \mid q \in Q_n\} \cup \{(q, b, q \cdot b) \mid q \in Q_n\}$. Thus every word u of A^* defines a permutation on Q_n (also denoted u). Conversely, since the symmetric group $\mathfrak{S}_n$ on Q_n is generated by a and b, every permutation on Q_n is represented by some word in A^*. Set, for $0 \leq k \leq n-1$, $\mathcal{A}_{n,k} = (Q_n, A, E, k, k)$ and let $L_{n,k} = |\mathcal{A}_{n,k}|$. Thus $L_{n,k}$ is the set of all words that represent a permutation on Q_n having k as a fixpoint. Clearly, every $\mathcal{A}_{n,k}$ is an n-state reversible automaton. It follows that the language $L_n = \cup_{0 \leq k \leq n-1} L_{n,k}$ is accepted by an n^2-state reversible automaton (the disjoint union of the $\mathcal{A}_{n,k}$'s).

It is easy to construct an automaton with $n!$ states accepting L_n: let $\mathfrak{S}_n$ be the set of states, let a and b act on $\mathfrak{S}_n$ by right multiplication and take the identity as initial state and the set of all permutations having at least one fixpoint as set of final states. The resulting automaton accepts L_n.

We now show that the minimal automaton of L_n has precisely $n!$ states. Let $\mathcal{A} = (Q, A, E, q_0, F)$ be this minimal automaton. To avoid any confusion, we shall denote $q \cdot_{\mathcal{A}} u$ the action of a word u on a state q of $\mathcal{A}$. We claim that if two words u and v define distinct permutations on Q_n, then $q_0 \cdot_{\mathcal{A}} u \neq q_0 \cdot_{\mathcal{A}} v$. Since a and b generate $\mathfrak{S}_n$, this will imply that $\mathcal{A}$ contains at least $n!$ states. Assume, by contradiction, that $q_0 \cdot u = q_0 \cdot v$. Since u and v are distinct permutations, there exists a state $q \in Q_n$ such that $q \cdot u \neq q \cdot v$. Let w be a word representing the permutation defined as follows

$$s \cdot w = (s + q \cdot v - q \cdot u)v^{-1} \quad \text{for every } s \in Q_n$$

where arithmetic operations are calculated modulo n and v^{-1} is the inverse permutation of v. We have by definition

$$q \cdot uw = (q \cdot u + q \cdot v - q \cdot u)v^{-1} = (q \cdot v)v^{-1} = q$$

and thus $uw \in L_q$ and $uw \in L$. On the other hand, we have for every $s \in Q$,

$$s \cdot vw = (s \cdot v + q \cdot v - q \cdot u)v^{-1} \neq s$$

for otherwise, $s \cdot v + q \cdot v - q \cdot u = s \cdot v$ and $q \cdot v = q \cdot u$. Therefore the permutation represented by vw has no fixpoint and thus $vw \notin L$. It follows that $q_0 \cdot_{\mathcal{A}} u \neq q_0 \cdot_{\mathcal{A}} v$, proving the claim and the proposition. □

In the opposite direction, the construction of a reversible automaton given in this article, although effective, would give an enormous upper bound for $R(n)$, and we don't have so far any reasonable bound to propose.

Acknowledgement

Theorem 3.2 and condition (2) of Theorem 7.1 were suggested by Christophe Reutenauer. Several computations were made possible thanks to the software AUTOMATE developped by Jean-Marc Champarnaud [8].

References

[1] D. Angluin, Inference of reversible languages, *Journal of the Association for Computing Machinery*,**29,** (1982) 741–765.

[2] C.J. Ash, Finite semigroups with commuting idempotents, *J. Austral. Math. Soc. (Series A)* **43**, (1987) 81–90.

[3] C.J. Ash, Inevitable sequences and a proof of the type II conjecture, in *Proceedings of the Monash Conference on Semigroup Theory*, World Scientific, Singapore, (1991) 31–42.

[4] C.J. Ash, Inevitable Graphs: A proof of the type II conjecture and some related decision procedures, *Int. Jour. Alg. and Comp.* **1** (1991) 127–146.

[5] J. Berstel, *Transductions and Context Free Languages*, Teubner Verlag, 1979.

[6] J.C. Birget, Intersection and union of regular languages, and state complexity, to appear.

[7] J.C. Birget, S.W. Margolis and J. Rhodes, Finite semigroups whose idempotents commute or form a subsemigroup, *in Semigroups and Their Applications*, edited by S.M. Goberstein and P.M. Higgins, Reidel, Dordrecht, 1987, 25–35.

[8] J.M. Champarnaud and G. Hansel, A computing package for automata and finite semigroups, *Journal of Symbolic Computation* **12**, 1991, 197–220.

[9] S. Eilenberg, *Automata, Languages and Machines*, Academic Press, New York, Vol. A, 1974; Vol B, 1976.

[10] M. Hall Jr., A topology for free groups and related groups, *Ann. of Maths* **52**, (1950) 127–139.

[11] T.E. Hall, Biprefix codes, inverse semigroups and syntactic monoids of injective automata, *Theoretical Computer Science.*

[12] K. Henckell, S.W. Margolis, J.E. Pin and J. Rhodes, Ash's type II theorem, profinite topology and Malcev products, to appear in *Int. Jour. Alg. and Comp.*.

[13] G. Lallement, *Semigroups and Combinatorial Applications*, Wiley, New York, 1979.

[14] M. Lothaire, *Combinatorics on words*, Encyclopedia of Mathematics **17**, Addison-Wesley, Reading, MA, 1983.

[15] R. McNaughton, The loop complexity of pure-group events. *Inf. Control* **11**, (1967) 167–176.

[16] S.W. Margolis, Consequences of Ash's proof of the Rhodes Type II Conjecture, in *Proceedings of the Monash Conference on Semigroup Theory*, World Scientific, Singapore, (1991) 180–205.

[17] S.W. Margolis and J.E. Pin, Languages and inverse semigroups, *11th ICALP, Lecture Notes in Computer Science* **199**, Springer, Berlin (1985) 285–299.

[18] S.W. Margolis and J.E. Pin, Inverse semigroups and varieties of finite semigroups, *Journal of Algebra* **110** (1987) 306–323.

[19] S.W. Margolis and J.E. Pin, New results on the conjecture of Rhodes and on the topological conjecture, to appear in *J. Pure and Applied Algebra.*

[20] J.E. Pin, Finite group topology and p-adic topology for free monoids. *12th ICALP, Lecture Notes in Computer Science 199*, Springer, Berlin, 1985, 285–299.

[21] J.E. Pin, *Variétés de langages formels*, 160 p., Masson, Paris (1984). *Varieties of formal languages*, 138 p., North Oxford Academic (London), 1986 and Plenum (New York), 1986.

[22] J.E. Pin, *On the languages recognized by finite reversible automata,* 14th ICALP, Lecture Notes in Computer Science 267 Springer, Berlin, (1987) 237–249.

[23] J.E. Pin, A topological approach to a conjecture of Rhodes, *Bulletin of the Australian Mathematical Society* **38** (1988) 421–431.

[24] J.E. Pin, On a conjecture of Rhodes, *Semigroup Forum* **39** (1989) 1–15.

[25] J.E. Pin, Topologies for the free monoid, *Journal of Algebra* **137** (1991) 297–337.

[26] J.E. Pin and Ch. Reutenauer, A conjecture on the Hall topology for the free group, to appear in the *Notices of the London Math. Society.*

[27] Ch. Reutenauer, Une topologie du monoïde libre, *Semigroup Forum* **18**, (1979) 33–49.

[28] Ch. Reutenauer, Sur mon article "Une topologie du monoïde libre", *Semigroup Forum* **22**, (1981) 93–95.

[29] L. Ribes and P.A. Zalesskii, On the profinite topology on a free group, to appear.