On the construction of reversible automata for reversible languages

Sylvain Lombardy

Laboratoire Traitement et Communication de l'Information
Ecole Nationale Supérieure des Télécommunications
46, rue Barrault 75634 Paris Cedex 13, France
{lombardy}@enst.fr

Abstract. Reversible languages occur in many different domains. Although the decision for the membership of reversible languages was solved in 1992 by Pin, an effective construction of a reversible automaton for a reversible language was still unknown. We give in this paper a method to compute a reversible automaton from the minimal automaton of a reversible language. With this intention, we use the universal automaton of the language that can be obtained from the minimal automaton and that contains an equivalent automaton which is quasi-reversible. This quasi-reversible automaton has nearly the same properties as a reversible one and can easily be turned into a reversible automaton.

Key words : Finite automata, reversible languages, reversible automata, universal automata
Track : A

Introduction

Reversible languages are a class of rational languages that stands at the junction of several domains. P. Silva [12] has considered them in the study of inversive semigroups. J.-E. Pin [9] and P.-C. Héam [5] have given some topological properties of reversible languages. They constitute also a positive variety of rational languages and are in correpondence with a variety of ordered monoids: $\mathbf{E}^{-}_{\mathbf{com}}$ [10]. Among rational languages, they are a natural generalization to the notion of group languages. It is the reason why, in order to extend results proved on group languages, studying this class can prove to be wise. C. Nicaud [8] has studied average complexity of some operations on reversible automata. S. Lombardy and J. Sakarovitch [6] have given a new effective proof to compute their star height.

J.-E. Pin has given some characterizations of reversible languages in [9]. He gives also an algorithm to decide whether a rational language is reversible. Unfortunately, his paper does not provide any effective method to build a reversible automaton.

We present here such a construction from the minimal automaton of a reversible language. We use a canonical automaton, called *universal automaton*, attached to every rational language. On the one hand, we give an effective new method to compute this automaton from the minimal automaton of the language; on the other hand, we prove that this automaton contains a *quasi-reversible* automaton that accepts the language. As it is very easy to turn quasi-reversible automata into reversible ones, this provides a construction for reversible automata.

In a first part, we give some basic definitions about automata, that will be used in the course of this paper. In the second part, we define reversible languages. We recall the characterization given by J.-E. Pin[9]. Then, we introduce quasi-reversible automata, that accept reversible languages, are smaller, and give easily reversible automata.

In a third part, we present the universal automaton of a rational language, that is based on an idea of J.H. Conway [3]. This automaton has been studied by A. Arnold, A. Dicky, and M. Nivat [1], O. Matz and A. Potthoff [7] and J. Sakarovitch [11]. The main property of this finite automaton is that there exists a morphism from any automaton that accepts the same language, into the universal automaton. We give an effective construction of it.

The fourth part is devoted to the study of universal automata of reversible languages. We show that their strongly connected components are reversible. This result is the key of the theorem of the fifth part which claims that the universal automaton of a reversible language contains a quasi-reversible automaton that accepts this language.

1 Definitions

We denote by A^* the free monoid generated by a set A. Elements of A^* are words, the identity of this monoid is the empty word 1_{A^*}.

Definition 1. *We denote a (finite) automaton by a 5-tuple $\langle Q, A, E, I, T \rangle$, where Q is a finite set of states, A is a finite set of letters, E, the set of transitions, is a subset of $Q \times A \times Q$, and I (resp. T), the set of initial states (resp. terminal states), is a subset of Q.*

Definition 2. *Let $\mathcal{A}$ be an automaton and p a state of $\mathcal{A}$. The* ***past*** *of p in $\mathcal{A}$ is the set of words that label a path from an initial state of $\mathcal{A}$ to p. The* ***future*** *of p in $\mathcal{A}$ is the set of words that label a path from p to a terminal state of $\mathcal{A}$. We respectively denote the past and the future of p in $\mathcal{A}$ by $\mathsf{Past}_{\mathcal{A}}(p)$ and $\mathsf{Fut}_{\mathcal{A}}(p)$.*

Definition 3. *Let $\mathcal{A} = \langle Q, A, E, I, T \rangle$ and $\mathcal{B} = \langle R, A, F, J, U \rangle$ be two automata. A mapping μ from Q into R is a* ***morphism*** *of automata if and only if:*

$$p \in I \Rightarrow p\mu \in J, \qquad p \in T \Rightarrow p\mu \in U, \qquad \text{and} \quad (p, a, q) \in E \Rightarrow (p\mu, a, q\mu) \in F.$$

Proposition 1. *Let μ be a morphism from an automaton $\mathcal{A}$ into an automaton $\mathcal{B}$. Then, for every state p of $\mathcal{A}$,*

$$\mathsf{Past}_{\mathcal{A}}(p) \subseteq \mathsf{Past}_{\mathcal{B}}(p\mu), \quad \mathsf{Fut}_{\mathcal{A}}(p) \subseteq \mathsf{Fut}_{\mathcal{B}}(p\mu).$$

The proof is by induction on the length of words. □

Definition 4. *Let $\mathcal{A} = \langle Q, A, E, I, T \rangle$ be an automaton. For every state p in Q, for every letter a, we denote $p \cdot a$ the set $\{q \in Q \mid (p, a, q) \in E\}$. With the convention $p \cdot 1_{A^*} = \{p\}$, we can extend this definition to subsets of Q and to words:*

$$\forall u = u'a \in A^+, p \cdot u = (p \cdot u') \cdot a, \qquad \forall P \subseteq Q, P \cdot u = \bigcup_{p \in P} p \cdot u.$$

Symetrically, for every subset P of Q, for every word u,

$$u \cdot P = \{p \in Q \mid p \cdot u \cap P \neq \emptyset\}.$$

There is a path labelled by a word u between two states p and q if and only if q belongs to $p \cdot u$.

Definition 5. *A* ***strongly connected component*** *(SCC for short) of an automaton is a maximal subautomaton such that there is a path between every pair of states.*

2 Reversible languages : definition and realization

Definition 6. *An automaton $\mathcal{A} = \langle Q, A, E, I, T \rangle$ is* ***reversible*** *if, for every state p in Q, for every letter a in A there exists at most one transition in E that comes from p (resp. goes to p) with label a.*
A rational language is reversible if there exists a reversible automaton that accepts it.

Remark 1. A reversible automaton may have several initial or final states. As a consequence, the minimal automaton of a reversible language may be not reversible.

Theorem 1. *[9] There is a polynomial time algorithm for testing whether the language accepted by an automaton can be accepted by a reversible automaton.*

We give now a new representation for reversible languages.

Definition 7. *An automaton $\mathcal{A} = \langle Q, A, E, I, T \rangle$ is* ***quasi-reversible*** *if, in case two transitions of E have the same label and arrive in or come from the same state, none of them belongs to any SCC.*

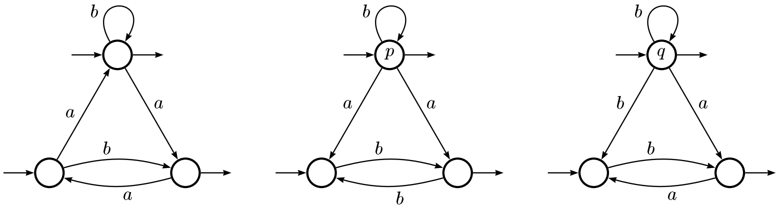

Fig. 1. A reversible, a quasi-reversible and a non-reversible automata.

Example 1. Among the three automata of Figure 1, the first automaton is reversible. The second one is quasi-reversible: both transitions that come from p with label a do not belong to any SCC. The last one is not quasi-reversible: one of both transitions that come from q with label b does belong to a SCC (it is a loop).

The following proposition proves that languages that are accepted by quasi-reversible automata are reversible.

Proposition 2. *If $\mathcal{A}$ is a quasi-reversible automaton, one can build a reversible automaton that is equivalent to $\mathcal{A}$.*

Proof. We call sensitive pair a pair of transitions that is in contradiction with reversibility. The proof is by induction on the number of sensitive pair in $\mathcal{A}$. If there is no such a pair, then $\mathcal{A}$ is reversible. If there are n sensitive pairs, let $\{e_1, e_2\}$ be one of them. As neither e_1 nor e_2 belongs to any SCC, there is no path that contains these two transitions. One creates two copies $\mathcal{A}_1$ and $\mathcal{A}_2$ of $\mathcal{A}$ and removes e_1 from $\mathcal{A}_1$ and e_2 from $\mathcal{A}_2$. The number of sensitive pair is smaller in each of this two quasi-reversible automata. Thus, by induction, they can be turned into reversible automata. □

Remark 2. One also could transform the quasi-reversible automaton into a set of reversible "string-like"-automata, as we shall see later.

Quasi-reversible automaton can be much smaller than reversible ones. P.-C. Héam [4] has given the following example: the language $(aa + ab + bb)^{2n}$ is finite and is therefore reversible. The minimal automaton of this language is quasi-reversible and has $6n+1$ states, whereas the smallest reversible automaton that accepts this language has at least $(3\sqrt{2}/4)^n$ states.

3 Universal automaton

The universal automaton of a language has been defined by Conway [3] as a "factor matrix". A complete study of properties of this automaton can be found in [11]. We recall here the definition and some basic properties of this automaton. Lastly, we introduce a new construction for the universal automaton.

Definition 8. *Let $\mathcal{L}$ be a rational language of A^*. A **factorization** of $\mathcal{L}$ in A^* is a maximal couple (with respect to the inclusion) of languages (L,R) such*

that $L.R \subseteq \mathcal{L}$.
The ***universal automaton*** *of $\mathcal{L}$ is $\mathcal{U}_{\mathcal{L}} = \langle Q, A, E, I, T \rangle$, where Q is the set of factorizations of $\mathcal{L}$, and*

$$\begin{aligned}
I =&\{(L,R) \in Q \mid 1_{A^*} \in L\},\\
T =&\{(L,R) \in Q \mid \mathcal{L} \subseteq L\},\\
E =&\{((L,R),a,(L',R')) \in Q \times a \times Q \mid L.a \subseteq L'\}.
\end{aligned}$$

Remark 3. As factorizations are maximal, the condition $1_{A^*} \in L$ is equivalent to $\mathcal{L} \subseteq R$ and $\mathcal{L} \subseteq L$ is equivalent to $1_{A^*} \in R$. For the same reason, $L.a \subseteq L'$ is equivalent to $a.R' \subseteq R$ and to $L.a.R' \subseteq \mathcal{L}$.

Proposition 3. *The universal automaton of a rational language is a finite automaton.*

Factorizations of a rational language are recognized by the syntactic monoid of this language. Thus there is a finite number of factorizations. □

Example 2. Let $\mathcal{L}_1 = b^*ab((a+b)b^*+1)+b^*(a+1)$. Figure 2 shows the minimal automaton of this language.

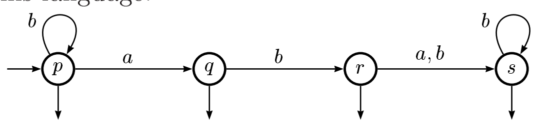

Fig. 2. Minimal automaton of $\mathcal{L}_1$.

By computing the syntactic monoid, we can get factorizations of $\mathcal{L}_1$ in $(a+b)^*$:

$$\begin{gathered}
(b^*, \mathcal{L}_1), \qquad (b^*(a+1), (ba+1)b^*)\\
(b^*(ab+1), (a+1)b^*), \qquad (\mathcal{L}_1, b^*).
\end{gathered}$$

Hence, the universal automaton of $\mathcal{L}_1$ has four states. They are all initial, because 1_{A^*} belongs to the left factor of every factorization and they are all terminal because the left factor of every factorization is a subset of $\mathcal{L}_1$. We show at the end of this section that we can build the universal automaton without computing these factorizations.

Proposition 4. *Let (L,R) be a state of a universal automaton $\mathcal{U}_{\mathcal{L}}$. Then:*

$$\mathsf{Past}_{\mathcal{U}_{\mathcal{L}}}((L,R)) = L \quad \mathsf{Fut}_{\mathcal{U}_{\mathcal{L}}}((L,R)) = R.$$

The proof is by induction on the length of words. □

Corollary 1. *The universal automaton of a rational language $\mathcal{L}$ accepts $\mathcal{L}$.*

Proof. For every terminal state (L,R) of the universal automaton, $\mathsf{Past}_{\mathcal{U}_{\mathcal{L}}}(L)$ is a subset of $\mathcal{L}$. Thus, the language accepted by $\mathcal{U}_{\mathcal{L}}$ is a subset of $\mathcal{L}$.
Let $L = \{u \in A^* \mid u.\mathcal{L} \subseteq \mathcal{L}\}$. Then, $(L,\mathcal{L})$ is a factorization which is an initial state of $\mathcal{U}_{\mathcal{L}}$ and is future is $\mathcal{L}$. Hence, $\mathcal{U}_{\mathcal{L}}$ accepts $\mathcal{L}$. □

We are here especially interested in the following property of the universal automaton.

Proposition 5. *Let $\mathcal{L}$ be a rational language, and $\mathcal{A}$ be a trim automaton that accepts $\mathcal{L}$. Then there exists a morphism from $\mathcal{A}$ into the universal automaton of $\mathcal{L}$.*

We build a mapping μ from states of $\mathcal{A}$ into factorizations of $\mathcal{L}$: $p\mu = (L_p, R_p)$, with:

$$R_p = \{v \in A^* \mid \mathsf{Past}_{\mathcal{A}}(p).v \subseteq \mathcal{L}\}, \qquad L_p = \{u \in A^* \mid u.R_p \subseteq \mathcal{L}\}.$$

We check then that μ is a morphism between automata. □

The following proposition gives an effective method to build the universal automaton of a rational language from its minimal automaton.

Proposition 6. *Let $\mathcal{A} = \langle Q, A, E, \{i\}, T \rangle$ be the minimal automaton of $\mathcal{L}$ and P be the set of states of the codeterminized automaton of $\mathcal{A}$:*

$$P = \{X \subseteq Q \mid \exists u \in A^*, X = u \cdot T\}.$$

Let $P_\cap$ be the closure of P under intersection (without the empty set):

$$X, Y \in P_\cap, X \cap Y \neq \emptyset \Longrightarrow X \cap Y \in P_\cap.$$

Then, the universal automaton $\mathcal{U}_{\mathcal{L}}$ is isomorphic to $\langle P_\cap, A, F, J, U \rangle$, with:

$$\begin{aligned} J =& \{X \mid i \in X\} \\ U =& \{X \mid X \subseteq T\} \\ F =& \{(X, a, Y) \mid X \cdot a \subseteq Y \text{ and } \forall p \in X, p \cdot a \neq \emptyset\}. \end{aligned}$$

Proof. Let φ be the mapping from $P_\cap$ into $A^* \times A^*$ defined by $X\varphi = (L_X, R_X)$, with:

$$L_X = \bigcup_{p \in X} \mathsf{Past}_{\mathcal{A}}(p), \qquad R_X = \bigcap_{p \in X} \mathsf{Fut}_{\mathcal{A}}(p).$$

For every X in $P_\cap$, $X = \bigcap_{u \in R_X} u \cdot T$. Hence, (L_X, R_X) is a factorization, and, for every factorization (L, R), the set $X = \bigcap_{u \in R} u \cdot T$ belongs to $P_\cap$ and $X\varphi = (L, R)$. Thus φ is a one-to-one mapping between $P_\cap$ and factorizations of $\mathcal{L}$. Then, we check that φ is an isomorphism of automata. □

Remark 4. The minimality of $\mathcal{A}$ is not required in this construction. WE only assume that it is deterministic. We can show that the cardinality of $P_\cap$ is not affected by the minimality of $\mathcal{A}$.

Example 3. The minimal automaton of $\mathcal{L}_1$ is shown on Figure 2. Figure 3 shows its codeterminized automaton. The set $\{\{p\}, \{p, q\}, \{p, r\}, \{p, q, r, s\}\}$ is closed

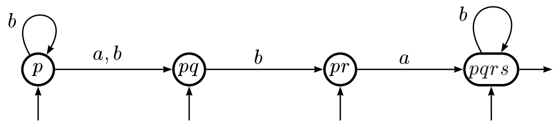

Fig. 3. The codeterminized automaton of the minimal automaton of $\mathcal{L}_1$.

under intersection. Hence, the automaton defined Proposition 6 (and isomorphic to the universal automaton) has the following transition table:

	$\{p\}$	$\{p,q\}$	$\{p,r\}$	$\{p,q,r,s\}$
$\{p\}$	b	a,b	b	a,b
$\{p,q\}$			b	b
$\{p,r\}$				a,b
$\{p,q,r,s\}$				b

All the states are initial and terminal. The universal automaton of $\mathcal{L}_1$ is drawn on Figure 4.

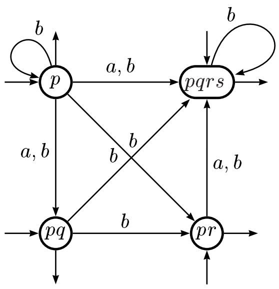

Fig. 4. Universal automaton of $\mathcal{L}_1$.

4 Strongly connected components of the universal automaton of a reversible language

We show in this part that every SCC of the universal automaton of a reversible language is reversible. This result is an important step to prove that there is a quasi-reversible automaton that accepts the language in the universal automaton. We define first the subset expansion of an automaton. We show that the universal automaton is one of its subautomata and that the SCC of the subset expansion of a reversible automaton are reversible.

Definition 9. *Let $\mathcal{A} = \langle Q, A, E, I, T \rangle$ be an automaton that accepts $\mathcal{L}$. Let S be the set of antichains of $\mathcal{P}(Q)$ (the set of subsets of Q, partially ordered by inclusion). The* ***subset expansion*** *of $\mathcal{A}$ is the automaton $\mathcal{V}_{\mathcal{A}} = \langle S, A, F, J, U \rangle$, with:*

$$\begin{aligned} J =& \{X \in S \mid \exists Y \in X, Y \subseteq I\}, \\ U =& \{X \in S \mid \forall Y \in X, Y \cap T \neq \emptyset\}, \\ F =& \{(X, a, X') \mid \forall Y \in X, \exists Y' \in X', Y' \subseteq Y \cdot a\}. \end{aligned}$$

Remark 5. If $\mathcal{A}$ is a deterministic automaton, the subset expansion of $\mathcal{A}$ is exactly the *subset automaton of order* 0 introduced by R. Cohen and J.A. Brzozoski [2]. In this case, the construction is the same as Proposition 6, excepted that the set of states is isomorphic to $\mathcal{P}(Q) \setminus \{\emptyset\}$ instead of being isomorphic to $P_\cap$ (which is smaller) defined in the proposition.

Lemma 1. *The subset expansion of an automaton $\mathcal{A}$ accepts the same language as $\mathcal{A}$.*

Proof. Let I (*resp.* T) be the set of initial (*resp.* terminal) states of $\mathcal{A}$. We prove by induction on the length of words, that there is a path labelled by u between X and X' if and only if:

$$\forall Y \in X, \exists Y' \in X', Y' \subseteq Y \cdot u$$

If a word u is accepted by $\mathcal{V}_\mathcal{A}$, there exists a path labelled by u between an initial state X and a terminal state X'. There exists Y in X such that $Y \subseteq I$. Hence, there exists Y' in X' such that $Y' \subseteq Y \cdot u \subseteq I \cdot u$. As X' is terminal, $Y' \cap T \neq \emptyset$ and $I \cdot u \cap T \neq \emptyset$. Therefore, the word u is accepted by $\mathcal{A}$.
Conversely, if u is accepted by $\mathcal{A}$, $I \cdot u \cap T \neq \emptyset$, hence, there is a path labelled by u between the initial state $\{I\}$ and the terminal state $\{\{p\} \mid p \in T\}$. □

Proposition 7. *Let $\mathcal{A}$ be an automaton that accepts $\mathcal{L}$. The universal automaton of $\mathcal{L}$ is a subautomaton of the subset expansion of $\mathcal{A}$.*

We build a mapping from states of $\mathcal{U}_\mathcal{L}$ into states of $\mathcal{V}_\mathcal{A}$:

$$(L, R) \longmapsto \mathsf{min}\{I \cdot u \mid u \in L\},$$

where I is the set of initial states of $\mathcal{A}$. We show then that this is an injective morphism of automata. □

Proposition 8. *Let $\mathcal{A}$ be a reversible automaton. The SCC of the subset expansion of $\mathcal{A}$ are reversible.*

Proof. Let $\mathcal{V}_\mathcal{A}$ be the subset expansion of $\mathcal{A}$. We show that in every SCC of $\mathcal{V}_\mathcal{A}$, every word induces a one-to-one mapping between sets that characterize states. Let p and q two states of the same SCC. There exist two words u and v that respectively label paths from p to q and from q to p. Let X_p and X_q be the corresponding elements of S (*cf.* Definition 9).
We show by induction on the cardinal k of the elements of X_p (*resp.* X_q) that u (*resp.* v) induces a bijection from X_p onto X_q (*resp.* from X_q onto X_p) between elements of same cardinal.

There is no element of cardinal zero, thus the base of the induction is trivial. Let Y be an element of X_p of cardinal k.

i) As $\mathcal{A}$ has a deterministic transition system, the cardinal of $Y \cdot u$ is smaller than or equal to k. If it is smaller, by induction hypothesis, there exists Y' in X_p with the same cardinal as $Y \cdot u$ such that $Y' \cdot u = Y \cdot u$. As $\mathcal{A}$ is codeterministic,

$Y' \subset Y$, which is in contradiction with the fact that X is an antichain. Therefore, $Y \cdot u$ has the same cardinal as Y and u is a bijection from elements of Y onto elements of $Y \cdot u$.

ii) If $Y \cdot u$ does not belong to X_q, there is a smaller element Z in X_q in bijection with an element Y' in X_p such that $Y' \cdot u = Z \subset Y \cdot u$. Once more, it implies $Y' \subset Y$, which is impossible, because X is an antichain.

iii) If $Y \cdot u = Y' \cdot u$, as Y is in bijection with $Y \cdot u$, it means that $Y \subseteq Y'$. If Y' belongs to X_p, $Y = Y'$. Thus u induces an injective function from elements of X_p of cardinal k into elements of X_q of same cardinal. As v induces a similar injection from X_q into X_p, these functions are one-to-one.

Therefore, from a state p, given a word u, there is at most one state in the same SCC, that can be reached by a path labelled by u. Thus, SCC are deterministic and, symetrically, codeterministic. □

Corollary 2. *The SCC of the universal automaton of a reversible language are reversible.*

5 Universal automaton of a reversible language

We prove in this part that the universal automaton of a reversible language contains a quasi-reversible subautomaton that accepts the language. For this purpose, we show that, for every reversible language, there exists a reversible automaton such that its image, in the universal automaton of the language, is quasi-reversible. This reversible automaton has a particular form. It is a union of "string-like"-automata:

Definition 10. *An automaton is a* ***string-like automaton*** *(SL-automaton for short) if*

i) it has one initial and one terminal state,
ii) there is at most one transition that comes from (resp. goes to) every SCC,
iii) it is connected,
iv) it is trim.

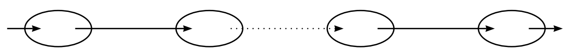

Fig. 5. Shape of a SL-automaton.

The general shape of a SL-automaton is presented Figure 5, where ellipses are SCC (that may be a single state).

Proposition 9. *Every reversible automaton is equivalent to a finite union of reversible SL-automata.*

The construction is illustrated on Figure 6. If the automaton is reversible (or merely quasi-reversible) every SL-automaton is reversible. □

Definition 11. *A union of SL-automata is minimal if one can not erase any SL-automaton without changing the language accepted by the union.*

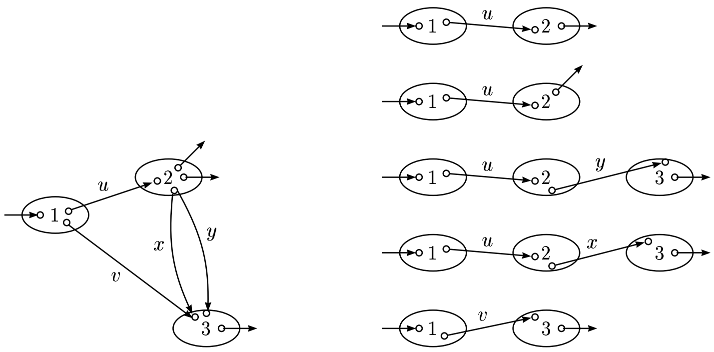

Fig. 6. From an automaton to a set of SL-automata.

In other words, every SL-automaton of a minimal union accepts a word which is not accepted by the other components.

Proposition 10. *Let $\mathcal{L}$ be a reversible language and $\mathcal{A}$ an automaton that accepts $\mathcal{L}$ and that is a minimal finite union of reversible SL-automata. Let φ be a morphism from $\mathcal{A}$ into $\mathcal{U}_\mathcal{L}$. Then, the image of every SCC of $\mathcal{A}$ is a SCC of $\mathcal{U}_\mathcal{L}$.*

Proof. We show that, if there is a SCC $\mathcal{S}$ for which it is not true, every word accepted by the SL-automaton $\mathcal{R}$, that contains $\mathcal{S}$, is accepted by another SL-automaton of $\mathcal{A}$, which is in contradiction with the minimality of the set of SL-automata.

Let φ be a morphism from $\mathcal{A}$ into $\mathcal{U}_\mathcal{L}$. Let w be a word accepted only by $\mathcal{R}$ in $\mathcal{A}$. As $\mathcal{R}$ is string-like, in $\mathcal{R}$, every path crosses every SCC. Thus, there exists a state r in $\mathcal{S}$ that belongs to the path labelled by w. One can write $w = u.v$, where u (*resp.* v) is in $\mathsf{Past}_\mathcal{R}(r)$ (*resp.* in $\mathsf{Fut}_\mathcal{R}(r)$).

As $\mathcal{S}\varphi$ is not a SCC, there exists a transition e in the SCC that contains $\mathcal{S}\varphi$ but that does not belong to $\mathcal{S}\varphi$. Let x be a word labelling a loop around $r\varphi$ that passes through e. As u (*resp.* v) is in the past (*resp.* the future) of r and as $\mathcal{U}_\mathcal{L}$ recognizes $\mathcal{L}$, for every k in $\mathbb{N}$, the word $u.x^k.v$ is in $\mathcal{L}$. As $\mathcal{A}$ is reversible, there exists an integer n such that, for every state of $\mathcal{A}$, either x^n labels a loop around the state or there is no path starting in this state and labelled by x^n. The word $u.x^n.v$ is in $\mathcal{L}$. If it was accepted by $\mathcal{R}$, there would be a loop labelled by x^n around r, and, as $\mathcal{S}\varphi$ is deterministic, the transition e would belong to the image of this loop. Thus, this word is not accepted by $\mathcal{R}$ and there exists another SL-automaton of $\mathcal{A}$ that accepts it, and, as x^n labels a loop, that accepts $u.v$ too. Which is in contradiction with the assumption.

Therefore, the image of every SCC of every SL-automaton of $\mathcal{A}$ is a SCC. $\square$

Proposition 11. *Let $\mathcal{L}$ be a reversible language and $\mathcal{A}$ an automaton that accepts $\mathcal{L}$ and that is a minimal union of reversible SL-automata. Then the image of every SL-automaton of $\mathcal{A}$ in the universal automaton of $\mathcal{L}$, by any morphism, is reversible.*

Proof. If it is not true, let $\mathcal{R}$ be a SL-automaton of $\mathcal{A}$ the image of which contains two transitions with the same label that either arrive at the same state p or come from the same state p. We can assume without loss of generality that the first configuration occurs. Let t_1 and t_2 be these two transitions of $\mathcal{U}_\mathcal{L}$, that are respectively the images of s_1 and s_2. As the image of a SL-automaton is a SL-automaton and as SCC of $\mathcal{U}_\mathcal{L}$ are reversible, one of both transitions, t_1, for instance, is in a SCC $\mathcal{S}$ and the other one, t_2, is not. As $\mathcal{R}$ is a SL-automaton, there exists a path that starts at the end of s_2 and the last transition of which is s_1. The image of this path in $\mathcal{U}_\mathcal{L}$ is a loop labeled by x.
If $\mathcal{A}$ is a minimal union of SL-automata, there exists a word w that is accepted only by $\mathcal{R}$ in $\mathcal{A}$. As s_2 does not belong to any SCC, the path labelled by w in $\mathcal{R}$ passes through s_2. Let u be the label of the part of this path which ends with s_2 and v the label of the second part. Words u and v respectively belong to the past and the future of p in $\mathcal{U}_\mathcal{L}$. Thus $u.x^k.v$ belongs to $\mathcal{L}$ for every k, and, as in the proof of Proposition 10, with the same notations, we prove that there exists a SL-automaton of $\mathcal{A}$ that accepts $u.x^n.v$ (and that can therefore not be $\mathcal{R}$) and $u.v = w$, which is in contradiction with the assumption. □

Definition 12. *Let $\mathcal{U}_\mathcal{L}$ be the universal automaton of a reversible language $\mathcal{L}$. The maximum quasi-reversible subautomaton of $\mathcal{U}_\mathcal{L}$ is the largest subautomaton of $\mathcal{U}_\mathcal{L}$*

i) which is quasi-reversible,
ii) which has the same SCC as $\mathcal{U}_\mathcal{L}$.

This automaton exists because SCC of $\mathcal{U}_\mathcal{L}$ are reversible. It is obtained from $\mathcal{U}_\mathcal{L}$ by deleting transitions that do not belong to any SCC and that are in contradiction with the assumption of quasi-reversibility.

Theorem 2. *The maximum quasi-reversible subautomaton of the universal automaton of a reversible language accepts this language.*

Proof. Let $\mathcal{A}$ be a minimal finite union of reversible SL-automata that accepts $\mathcal{L}$. The image of every SL-automaton in $\mathcal{U}_\mathcal{L}$ by any morphism is a reversible automaton which covers SCC of $\mathcal{U}_\mathcal{L}$ that are intersected. Thus, the image of $\mathcal{A}$ is a subautomaton of the maximum quasi-reversible subautomaton of $\mathcal{U}_\mathcal{L}$. Therefore, this one accepts the language $\mathcal{L}$. □

Example 4. The language $\mathcal{L}_1$ is reversible. (It can be check on its syntactic monoid). Figure 7 shows the maximum quasi-reversible subautomaton of its universal automaton.

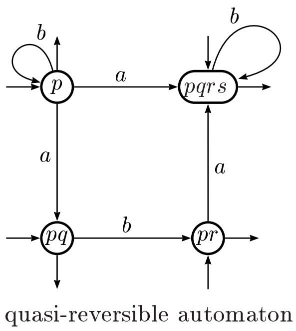

Fig. 7. A quasi-reversible automaton for $\mathcal{L}_1$.

6 Complexity of the algorithm

The size of the universal automaton may be exponential. Thus, the complexity of the algorithm that is described in this paper will be at least exponential.

The first step is the construction of the universal automaton. The computation of the codeterminized automaton and the closure of its states under intersection is polynomial in the size of the output (that may be exponential in the size of the input).

The second step is the computation of the maximal quasi-reversible subautomaton of the universal automaton. This requires the computation of the SCC, that is linear in the number of transitions. The number of transitions is quadratic in the number of states, but the square of an exponential number remains exponential. Then, the criterium of quasi-reversibility is local and the construction is polynomial in the size of the universal automaton.

Therefore, this algorithm is exponential in the size of the minimal automaton, but it is polynomial in the size of the output.

The transformation of a quasi-reversible automaton into a reversible one may be exponential, that makes the algorithm doubling exponential.

7 Acknoledgement

The author thanks Jacques Sakarovitch for his pertinent advices. He is grateful to Pierre-Cyrille Héam for invaluable conversations on this topic and others.

References

1. A. Arnold, A. Dicky, and M. Nivat, A note about minimal non-deterministic automata. *Bull. of E.A.T.C.S. 47* (1992), 166–169.
2. R. Cohen, et J. A. Brzozowski, General properties of star height of regular events, *J. Computer System Sci.* **4** (1970), 260–280.
3. J. H. Conway, *Regular algebra and finite machines*, Chapman and Hall, 1971.
4. P.-C. Héam, A lower bound for reversible automata, *Theoret. Informatics Appl.* **34** (2000), 331–341.
5. P.-C. Héam, Some topological properties of rational sets, *J. of Automata, Lang. and Comb.*, to appear.
6. S. Lombardy and J. Sakarovitch, Star height of reversible languages and universal automata, *accepted in Latin'02.*
7. O. Matz et A. Potthoff, Computing small nondeterministic finite automata. *proc. of TACAS'95, BRICS Notes Series* (1995), 74–88.
8. C. Nicaud, Etude du comportement en moyenne des automates finis et des langages rationnels, *Thèse de doctorat*, Université Paris 7, 2000.
9. J.-E. Pin, On reversible automata, *Proc. 1st LATIN Conf., (I. Simon, Ed.), Lecture Notes in Comput. Sci.* **583** (1992), 401–416.
10. J.-E. Pin, A variety theorem without complementation, *Russian mathematics* **39** (1995), 80–90.
11. J. Sakarovitch, *Eléments de théorie des automates*, Vuibert, to appear.
12. P.V. Silva, On free inverse monoid languages, *Theoret. Informatics and Appl.* **30** (1996), 349–378.