On the expressive power of temporal logic (*)

Joëlle Cohen, Dominique Perrin and Jean-Eric Pin

LITP, Paris, FRANCE

Abstract. We study the expressive power of linear propositional temporal logic interpreted on finite sequences or words. We first give a transparent proof of the fact that a formal language is expressible in this logic if and only if its syntactic semigroup is finite and aperiodic. This gives an effective algorithm to decide whether a given rational language is expressible. Our main result states a similar condition for the "restricted" temporal logic (RTL), obtained by discarding the "until" operator. A formal language is RTL-expressible if and only if its syntactic semigroup is finite and satisfies a certain simple algebraic condition. This leads to a polynomial time algorithm to check whether the formal language accepted by an n-state deterministic automaton is RTL-expressible.

Temporal logic is a particular case of modal logic. It was introduced by Pnueli [16] in connection with applications to the specification, development and verification of possibly parallel or non-deterministic processes. This logical language admits several variations, one of them being propositional linear temporal logic (PTL). It uses three connectives suggestively called "next", "eventually" and "until".

In this paper we are interested in the descriptive power of propositional linear temporal logic and of a restriction of temporal logic (RTL) obtained by considering only the operators "next" and "eventually". In both cases, we interpret temporal logic on finite words only. In this case, a temporal formula defines a set of words (that is, a formal language) and our problem is to determine precisely which formal languages can be specified in this way. In the case of PTL, the solution has been known for some time, as a consequence of a series of deep results. Indeed, Kamp [6] has shown that PTL is expressively equivalent to first-order logic when interpreted on words. Next, McNaughton [10] proved that a formal language is first-order definable if and only if it is star-free. Finally, star-free languages are characterized by a deep theorem of Schützenberger [17]: a rational (or regular) language is star-free if and only if its syntactic semigroup is group-free. Since the syntactic semigroup of a given rational language can be *effectively* computed, this provides an algorithm to determining whether a rational language is PTL-definable.

Various proofs of the equivalence between "first-order", "star-free" and "PTL-definable" have been announced or given in the literature [5,6,11,12] but all these proofs are rather involved. In this paper, we give a short and simple proof of the equivalence between star-free and PTL-definable, based on a weak version of the Krohn-Rhodes decomposition theorem for finite semigroups. Our proof was inspired by the work of [11], whose proof uses an interesting connection with Petri nets.

Our main result concerns the descriptive power of RTL. It was known [5,7] that

(*) Research on this paper was partially supported by PRC "Mathématiques et Informatique".

RTL is strictly less expressive than PTL, but an effective characterization of RTL-definable formal languages was still to be found. We show here that RTL-definable languages admit a syntactic characterization analogous to Schützenberger's theorem: a rational language is RTL-definable if and only if its syntactic semigroup is "locally $\mathcal{L}$-trivial". This provides a decision procedure to determine whether a formal language is RTL-definable. This algebraic characterization also leads to a polynomial time algorithm to check whether the formal language accepted by an n-state (complete) deterministic automaton is RTL-definable. We give another (non-effective) description of RTL-definable formal languages: these formal languages form the smallest boolean algebra of formal languages containing the languages aA^* and closed under the operations $L \to aL$ and $L \to A^*L$ for every letter a.

1. Semigroups and formal languages.

In this section, we briefly review some basic facts about finite semigroups and rational languages. All the definitions and results presented in this section are standard, and are reproduced for the convenience of the reader. More information on this subject can be found in [3,8,15]. For the most part, we follow the notations and terminology of Eilenberg [3]. In particular, if $\varphi : S \to T$ is a function from S into T, we denote by $s\varphi$ (instead of the usual $\varphi(s)$) the image of an element s of S by φ. We also use the term "rational language" instead of "regular language" for two reasons: first, the term "rational" has a much better mathematical foundation (rational languages are deeply connected with rational series), and second the term "regular" is also used in semigroup theory with a totally different meaning, and could be misleading in our context.

1.1. Semigroups.

A *semigroup* is a set S together with an associative multiplication. A *monoid* M is a semigroup that has an identity element, usually denoted by 1. The free monoid (resp. semigroup) on a set A is the set, usualy denoted A^* (resp. A^+) of all words (resp. non-empty words) over A, equipped with the concatenation of words as multiplication. Thus $A^* = A^+ \cup \{1\}$, where 1 is the empty word. Given two semigroups S and T, a semigroup morphism $\varphi : S \to T$ is a function from S into T such that, for every $s, s' \in S$,

$$(s\varphi)(s'\varphi) = (ss')\varphi.$$

All semigroups considered in this paper are finite except for free semigroups and free monoids. Therefore, we shall use in the sequel the term "semigroup" instead of "finite semigroup". An element e of a semigroup S is *idempotent* if $e^2 = e$. The set of idempotents of a semigroup S is denoted by $E(S)$. Every non-empty semigroup contains at least one idempotent. This is a particular case of the following well-known result:

Proposition 1.1. *For any semigroup S, there exists an integer $n \leq \mathrm{Card}(S)$ such that, for every $s \in S$, s^n is idempotent.*

The smallest integer n satisfying this property is called the *exponent* of S and is usually denoted $\omega(S)$ or simply ω. Thus s^ω is a convenient notation for the (unique) idempotent which is a power of s. For instance, if $x, y \in S$, $(x^\omega y^\omega)^\omega$ denotes the idempotent which is a power of ef, where e (resp. f) is the idempotent which is a power of x (resp. y). We shall frequently use this type of notation in the sequel. If S is a semigroup, the reverse semigroup S^r is the semigroup with underlying set S together with the operation $*$ defined by $s * t = ts$.

If S is a semigroup, we denote by S^1 the monoid equal to S if S is already a monoid, and otherwise equal to $S \cup \{1\}$, where 1 is a new identity element

We shall consider in particular three semigroups, denoted respectively U_1, U_2, and $B(1,2)$: U_1 is the semigroup $\{0,1\}$ with the multiplication given by $1.1 = 1$ and $0.1 = 1.0 = 0.0 = 0$, $B(1,2) = \{a,b\}$ with the multiplication given by $a.b = b.b = b$ and $a.a = b.a = a$, and $U_2 = B(1,2)^1 = \{1,a,b\}$.

The Green's relations $\mathcal{R}$ and $\mathcal{L}$ on a semigroup S are the equivalence relations defined as follows:

$$s \,\mathcal{R}\, t \text{ if and only if there exist } u,v \in S^1 \text{ such that } su = t \text{ and } tv = s,$$
$$s \,\mathcal{L}\, t \text{ if and only if there exist } u,v \in S^1 \text{ such that } us = t \text{ and } vt = s$$

A semigroup S is *$\mathcal{R}$-trivial* (respectively *$\mathcal{L}$-trivial*) if the relation $\mathcal{R}$ (respectively $\mathcal{L}$) is equality. For instance, U_1 is both $\mathcal{R}$-trivial and $\mathcal{L}$-trivial, and $B(1,2)$ and U_2 are $\mathcal{L}$ but not $\mathcal{R}$-trivial, since $a \,\mathcal{R}\, b$.

Given a semigroup S, and an idempotent e of S, the three subsets

$$eS = \{es \mid s \in S\}, \quad eSe = \{ese \mid s \in S\}, \quad Se = \{se \mid s \in S\}$$

are subsemigroups of S. The subsemigroup eSe is called the *local* semigroup associated with e. It is in fact a monoid, since e is clearly an identity of eSe. A semigroup S is said to have a property *locally* if for every idempotent e of S, the subsemigroup eSe has the property. In particular, a semigroup S is locally $\mathcal{R}$-trivial (respectively locally *$\mathcal{L}$-trivial*) if, for every idempotent e of S, eSe is $\mathcal{R}$-trivial (respectively $\mathcal{L}$-trivial). For instance, $B(1,2)$ is locally $\mathcal{R}$-trivial, but U_2 is not, since $1.U_2.1 = U_2$ is not $\mathcal{R}$-trivial.

Proposition 1.2. *Let S be a semigroup. Then*
(1) S is locally $\mathcal{R}$-trivial if and only if, for every $e \in E(S)$, Se is $\mathcal{R}$-trivial.
(2) S is locally $\mathcal{L}$-trivial if and only if, for every $e \in E(S)$, eS is $\mathcal{L}$-trivial.

Proof. Clearly, (2) is a dual version of (1). Let S be a locally $\mathcal{R}$-trivial semigroup. Let $e \in E(S)$, and suppose that $se \,\mathcal{R}\, te$ for some $s,t \in S$. Then there exist $ue, ve \in (Se)^1$ such that $seue = te$ and $teve = se$. Thus $s(eue)(eve) = se$. Furthermore $[(eue)(eve)]^\omega \,\mathcal{R}\, [(eue)(eve)]^\omega(eue)$ holds in eSe, and since eSe is $\mathcal{R}$-trivial, it follows

$$[(eue)(eve)]^\omega = [(eue)(eve)]^\omega(eue).$$

Therefore

$$se = s[(eue)(eve)]^\omega = s[(eue)(eve)]^\omega(eue) = s(eue) = te.$$

Conversely, assume that Se is $\mathcal{R}$-trivial. Then eSe, which is a subsemigroup of Se, is also $\mathcal{R}$-trivial. □

A semigroup S is *aperiodic* if for every $s \in S$, there exists an $n > 0$ such that $s^n = s^{n+1}$. For instance the three semigroups U_1, U_2, and $B(1,2)$ are aperiodic, but a non-trivial group is not aperiodic.

1.2. Transformation semigroups.

Let Q be a set, and let S be a semigroup. An *action* of S on Q is a function [(*)] from $Q \times S$ into Q, denoted $(q, s) \to q \cdot s$, such that, for every $q \in Q$ and every $s_1, s_2 \in S$,

$$(q \cdot s_1) \cdot s_2 = q \cdot (s_1 s_2).$$

Let $\mathcal{T}(Q)$ be the semigroup of all functions from Q into itself, with left-to-right composition of functions as the multiplication. Any action of S on Q defines a semigroup morphism $\rho : S \to \mathcal{T}(Q)$, given, for every $s \in S$, by

$$q \cdot (s\rho) = q \cdot s \quad \text{for every } q \in Q$$

The action of S on Q is *faithful* if ρ is injective, that is, if two elements of S having the same action on Q are equal. A *transformation semigroup* (*ts* for short) is a pair (Q, S), where Q is a set (the set of *states*) and S is a semigroup acting *faithfully* on Q.

Two natural examples of transformation semigroups are frequently used: first, every semigroup S defines a transformation semigroup (S^1, S), the action being simply the product in S. This transformation semigroup is usually denoted simply S, and the context suffices to decide whether one considers a semigroup or a transformation semigroup. The second example is the notion of *transformation semigroup of an automaton.* Let $\mathcal{A} = (Q, A, \cdot)$ be a (complete) deterministic automaton. By definition, every word w of A^+ defines a function $w\rho$ from Q into Q, given, for every $q \in Q$, by

$$q(w\rho) = q \cdot w$$

This defines a semigroup morphism $\rho : A^* \to \mathcal{T}(Q)$. The range of ρ is a subsemigroup of $\mathcal{T}(Q)$ denoted $S(\mathcal{A})$ and called the *semigroup of* A, and the transformation semigroup $TS(\mathcal{A}) = \big(Q, S(\mathcal{A})\big)$ is called the *transformation semigroup of* A. In practice, the notation $w\rho$ is almost always simplified to w, and the context makes clear whether one is considering w as a word or as an element of $\mathcal{T}(Q)$. For instance, if $\mathcal{A}$ is the automaton represented below

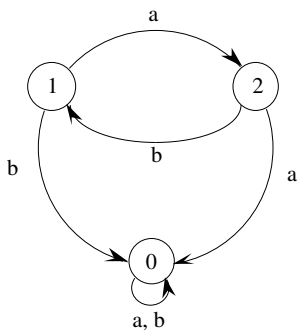


Figure 1.1.

[(*)] The definition of Eilenberg [3] allows partial functions, but we don't need this more general definition.

then $TS(\mathcal{A}) = (\{0,1,2\}, \{a, b, ab, ba, aa\})$ where the action of each element is represented in the following table:

	a	b	ab	ba	aa
0	0	0	0	0	0
1	2	0	1	0	0
2	0	1	0	2	0

We shall also use the transformation semigroup $\mathbf{2} = (\{1,2\}, B(1,2))$ where the action is given by the formulas $1 \cdot a = 2 \cdot a = 1$ and $1 \cdot b = 2 \cdot b = 2$.

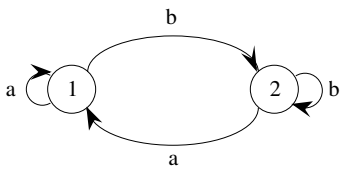


Figure 1.2.

A transformation semigroup (P, S) *divides* a transformation semigroup (Q, T) if there exists a surjective partial function $\varphi : Q \to P$, and, for every $s \in S$, there exists an element $\hat{s} \in T$ such that, for every $q \in Q$, $(q\varphi) \cdot s = (q \cdot \hat{s})\varphi$. For instance both $B(1,2)$ and U_1 divide U_2.

1.3. Formal languages.

Let A^+ be a free semigroup. The set A is called the *alphabet* and the elements of A are *letters*. The length of a word $w \in A^+$ is denoted by $|w|$. A subset of A^+ is called a *(formal) language*. *Rational languages* form the smallest class of languages containing letters and closed under union, concatenation and the plus operation ($L^+ = \bigcup_{n>0} L^n$). *Star-free languages* form the smallest class of languages containing letters and closed under boolean operations (union, intersection and complementation) and concatenation product.

The notion of the language recognized by an automaton can be easily adapted to transformation semigroups as follows: a transformation semigroup (Q, S) *recognizes* a language $L \subset A^+$ if there is a semigroup morphism $\eta : A^+ \to S$, a state $q_0 \in Q$ (the *initial state*), a set of states F (the *final states*) such that $L = \{u \in A^+ \mid q_0 \cdot (u\eta) \in F\}$. When the transformation semigroup is of the form $S = (S^1, S)$, there is a more convenient equivalent definition, that does not refer to transformation semigroups: a semigroup S recognizes a language $L \subset A^+$ if there is a morphism $\eta : A^+ \to S$, and a subset P of S, such that $L = P\eta^{-1}$. It is easy to see that if a language L is recognized by a transformation semigroup X, and if X divides a transformation semigroup Y, then Y also recognizes L.

For instance, if $a \in A$ and $B \subset A$, the languages A^*aA^*, A^*aB^* and A^*a are recognized by U_1, U_2 and $B(1,2)$, respectively. Conversely, we have the following lemma (see [15], chapter 2).

Lemma 1.3.

(1) If a language of A^+ is recognized by U_1, then it is a boolean combination of languages of the form A^*aA^* where $a \in A$.

(2) If a language of A^+ is recognized by U_2, then it is a boolean combination of languages of the form A^*aB^* where $a \in A$ and $B \subset A$.

(3) If a language of A^+ is recognized by $B(1,2)$, then it is a boolean combination of languages of the form A^*a where $a \in A$.

The *syntactic semigroup* of a language $L \subset A^+$, denoted $S(L)$, is the quotient of A^+ by the congruence $\sim_L$ defined by

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

The syntactic semigroup of a language L is the smallest semigroup that recognizes L. It is also the semigroup of the minimal automaton of L. As is well-known, a language is rational if and only if it can be recognized by a finite automaton. Since there are standard algorithms to compute the minimal automaton of a given rational language, this provides an algorithm to compute the syntactic semigroup of a rational language.

For star-free languages, we have the following important result, due to Schützenberger [17]. A proof can be found in [3,8,15,14].

Theorem 1.4. *Let L be a language. The following conditions are equivalent*

(1) L is star-free,

(2) L is recognized by an aperiodic semigroup,

(3) the syntactic semigroup of L is aperiodic.

1.4. Wreath product.

The *wreath product* of two transformation semigroups $X = (P, S)$ and $Y = (Q, T)$ is the transformation semigroup $X \circ Y = (P \times Q, S^Q \times T)$, with multiplication given by (*)

$$(f_1, t_1)(f_2, t_2) = (f, t_1 t_2), \quad \text{where, for every} q \in Q, \quad qf = (qf_1)(qt_1)f_2$$

and where the action of an element (f, t) of $S^Q \times T$ on a state (p, q) of $P \times Q$ is given by

$$(p, q) \cdot (f, t) = (p \cdot (qf), q \cdot t).$$

The wreath product is an associative operation on transformation semigroups. Aperiodic, $\mathcal{R}$-trivial and locally $\mathcal{R}$-trivial semigroups admit simple wreath-product decompositions using the three transformation semigroups U_1, U_2, and $\mathbf{2}$ defined in section 1.1 and 1.2 . For a proof, see [3] Vol. B or [20].

Theorem 1.5.

(1) A semigroup is $\mathcal{R}$-trivial if and only if it divides a wreath product of the form $U_1 \circ \cdots \circ U_1$.

(2) A semigroup is locally $\mathcal{R}$-trivial if and only if it divides a wreath product of the form $U_1 \circ \cdots \circ U_1 \circ \mathbf{2} \circ \cdots \circ \mathbf{2}$.

(3) A semigroup is aperiodic if and only if it divides a wreath product of the form $U_2 \circ \cdots \circ U_2$.

(*) S^Q denotes the set of all functions from Q to S. Thus if $f \in S^Q$ and $q \in Q$, qf is an element of S.

Wreath products are deeply related to sequential functions. Recall that a transducer $\mathcal{T} = (Q, A, B, q_0, ., *)$ is given by a finite set of states Q, an input alphabet A, an output alphabet B, an initial state q_0, a next-state function $Q \times A \to Q$, denoted $(q, a) \to q \cdot a$, and an output function $Q \times A \to B^+$, denoted $(q, a) \to q * a$. The next-state function is extended to a function $Q \times A^+ \to Q$ by setting $q \cdot (ua) = (q \cdot u) \cdot a$ for each $u \in A^*$ and $a \in A$. Similarly, the output function is extended to a function $Q \times A^+ \to B^+$ by setting $q * ua = (q * u)((q \cdot u) * a)$.

The function $\sigma : A^+ \to B^+$ defined by $u\sigma = q_0 * u$ is called the *sequential function* defined by $\mathcal{T}$. Then we can state

Proposition 1.6. [3] *Let $\sigma : A^+ \to B^+$ be a sequential function realized by a transducer $\mathcal{T} = (Q, A, B, q_0, \cdot, *)$ and let $S(\sigma)$ be the transformation semigroup of the automaton $(Q, A, \cdot)$. If a language $L \subset B^+$ is recognized by a semigroup S, then $L\sigma^{-1}$ is recognized by $S \circ S(\sigma)$.*

The following result is a first application of proposition 1.6 to a syntactic property of the operators $L \to LaA^*$ and $L \to La$ on languages.

Proposition 1.7. [3,22] *Let $L \subset A^+$ be a recognizable language. Then*
(1) $S(LaA^*)$ divides $U_1 \circ S(L)$,
(2) $S(La)$ divides $B(1,2) \circ S(L)$.

Proof. Let $\varphi : A^+ \to S = S(L)$ be the syntactic morphism of L. This morphism can be extended to a monoid morphism $\varphi : A^* \to S^1$. Put $P = L\varphi$, $B = S^1 \times A$, and let $\sigma : A^+ \to B^+$ be the sequential function defined by

$$(a_1 \cdots a_n)\sigma = (1\varphi, a_1) \cdots ((a_1 \cdots a_{n-1})\varphi, a_n).$$

Note that σ is realized by a transducer (that is, a deterministic automaton with output) with S^1 as the set of states and next-state and output functions defined by the following diagram.

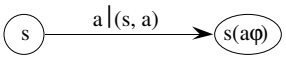


Figure 1.3.

In particular, the semigroup $S(\sigma)$ is equal to S. Put $C = P \times \{a\}$. Then C is a subset of B and we have

$$\begin{aligned}(B^*CB^*)\sigma^{-1} &= \{u \in A^+ \mid u\sigma \in B^*CB^*\} \\ &= \{a_1 \cdots a_n \in A^+ \mid \exists i \in \{1, \cdots, n-1\}\ ((a_1 \cdots a_i)\varphi a_{i+1}) \in C\} \\ &= \{a_1 \cdots a_n \in A^+ \mid \exists i \in \{1, \cdots, n-1\}\ a_1 \cdots a_i \in P\varphi^{-1} \text{ and } a_{i+1} = a\} \\ &= (P\varphi^{-1})aA^* = LaA^*.\end{aligned}$$

Therefore, by proposition 1.6, LaA^* is recognized by $S(B^*CB^*) \circ S(\sigma)$. Statement (1) follows, since $S(B^*CB^*) = U_1$. Similarly, we have

$$\begin{aligned}(B^*C)\sigma^{-1} &= \{u \in A^+ \mid u\sigma \in B^*C\} \\ &= \{a_1 \cdots a_n \in A^+ \mid ((a_1 \cdots a_{n-1})\varphi, a_n) \in B^*C\} \\ &= \{a_1 \cdots a_n \in A^+ \mid a_1 \cdots a_{n-1} \in P\varphi^{-1} \text{ and } a_n = a\} \\ &= (P\varphi^{-1})a = La.\end{aligned}$$

Therefore, by proposition 1.6, La is recognized by $S(B^*C) \circ S(\sigma)$. Statement (2) follows, since $S(B^*C) = B(1,2)$. □

Straubing's "wreath product principle" recalled below gives a description of the languages recognized by the wreath product of two transformation semigroups. Let $X = (P, S)$ and $Y = (Q, T)$ be two transformation semigroups, and let $Z = X \circ Y = (P \times Q, R)$, where $R = S^Q \times T$. Let L be a language of A^+ recognized by Z: then there exist an initial state $(p_0, q_0) \in P \times Q$, a set of final states F in $P \times Q$ and a morphism $\eta : A^+ \to R$ such that

$$L = \{u \in A^+ \mid (p_0, q_0) \cdot (u\eta) \in F\}.$$

The morphism η defines an action of A^+ on $P \times Q$ by setting $(p, q) \cdot a = (p, q)(a\eta)$.

Let π be the natural projection $\pi : R = S^Q \times T \to T$. Define a sequential function $\sigma : A^+ \to (Q \times A)^+$ by

$$(a_1 \cdots a_n)\sigma = (q_0, a_1)(q_0 \cdot (a_1\eta\pi), a_2) \cdots (q_0 \cdot (a_1 \cdots a_{n-1})\eta\pi, a_n).$$

We can now state

Proposition 1.8. (Wreath product principle [21]) *The language L is a finite union of languages of the form $U \cap V\sigma^{-1}$, where U is a language of A^+ recognized by Y and V is a language of $(Q \times A)^+$ recognized by X.*

Proposition 1.8, or some similar statement, together with theorem 1.5, has been used to prove theorem 1.4 [2,3,9].

1.5. Varieties of semigroups.

A variety of semigroups is a class of semigroups closed under taking subsemigroups, quotients and finite direct products (*). The following varieties will be used in this article:

A, the variety of aperiodic semigroups,
R, the variety of $\mathcal{R}$-trivial semigroups,
L, the variety of $\mathcal{L}$-trivial semigroups,
LR, the variety of locally $\mathcal{R}$-trivial semigroups,
LL, the variety of locally $\mathcal{L}$-trivial semigroups,

It is often convenient to define varities by identities. Let $u, v \in A^+$. Formally, a semigroup S satisfies the identity $u = v$ if and only if, for every semigroup morphism $\varphi : A^+ \to S$, $u\varphi = v\varphi$. For instance, a semigroup is commutative if and only if it satisfies the identity $xy = yx$. The next proposition gives identities defining the varieties **A**, **R**, **L**, **LR** and **LL**. In fact, there are not identities in the strict sense (**), since they involve the exponent ω, which depends on the semigroup S.

Proposition 1.9.

(1) A semigroup is aperiodic if and only if it satisfies the identity $x^\omega = x^{\omega+1}$,
(2) A semigroup is $\mathcal{R}$-trivial if and only if it satisfies the identity $(xy)^\omega x = (xy)^\omega$,
(3) A semigroup is $\mathcal{L}$-trivial if and only if it satisfies the identity $y(xy)^\omega = (xy)^\omega$,
(4) A semigroup is locally $\mathcal{R}$-trivial if and only if it satisfies the identity $(ux^\omega vx^\omega)^\omega ux^\omega = (ux^\omega vx^\omega)^\omega$, or, equivalently, the identity $(x^\omega ux^\omega vx^\omega)^\omega x^\omega ux^\omega = (x^\omega ux^\omega vx^\omega)^\omega$,
(5) A semigroup is locally $\mathcal{L}$-trivial if and only if it satisfies the identity $x^\omega v(x^\omega ux^\omega v)^\omega = (x^\omega ux^\omega v)^\omega$, or, equivalently, the identity $x^\omega v(x^\omega ux^\omega vx^\omega)^\omega = (x^\omega ux^\omega vx^\omega)^\omega$.

(*) The correct terminology should be "pseudovariety" to avoid a possible confusion with Birkhoff's varieties. However, we have preferred to avoid this rather awkward terminology.

(**) Again, the correct terminology should be "pseudoidentity".

A variety of semigroups $\mathbf{V}$ is closed under wreath product if, given two transformation semigroups $X = (P, S)$ and $Y = (Q, T)$ and their wreath product $(P \times Q, R)$, the conditions $S, T \in \mathbf{V}$ imply $R \in \mathbf{V}$. The next proposition is the "variety version" of theorem 1.5.

Proposition 1.10. [3,20]

(1) $\mathbf{R}$ *is the smallest variety of semigroups closed under wreath product containing* U_1.

(2) $\mathbf{LR}$ *is the smallest variety of semigroups closed under wreath product containing* U_1 *and* $B(1,2)$.

(3) $\mathbf{A}$ *is the smallest variety of semigroups closed under wreath product containing* U_2.

2. Propositional temporal logic.

Propositional temporal logic (PTL for short) on an alphabet A is defined as follows. The vocabulary consists of

(1) An atomic proposition p_a for each letter $a \in A$

(2) Connectives $\vee$, $\wedge$ and $\neg$.

(3) Temporal operators $\circ$ ("next"), $\diamond$ ("eventually") and $\mathcal{U}$ ("until").

and the formulas are constructed according to the rules

(1) For every $a \in A$, p_a is a formula,

(2) If φ and ψ are formulas, so are $\varphi \vee \psi$, $\varphi \wedge \psi$, $\neg\varphi$, $\circ\varphi$, $\diamond\varphi$, $\varphi \,\mathcal{U}\, \psi$.

Semantics are defined by induction on the formation rules. Given a word $w \in A^+$, and $n \in \{1, 2, ..., |w|\}$, we define the expression "w satisfies φ at the instant n" (denoted $(w, n) \models \varphi$) as follows

(1) $(w, n) \models p_a$ if the n-th letter of w is an a.

(2) $(w, n) \models \varphi \vee \psi$ (resp. $\varphi \wedge \psi$, $\neg\varphi$) if $(w, n) \models \varphi$ or $(w, n) \models \psi$ (resp. if $(w, n) \models \varphi$ and $(w, n) \models \psi$, if (w, n) does not satisfy φ).

(3) $(w, n) \models \circ\varphi$ if $(w, n+1)$ satisfies φ.

(4) $(w, n) \models \diamond\varphi$ if there exists m such that $n \leq m \leq |w|$ and $(w, m) \models \varphi$.

(5) $(w, n) \models \varphi \,\mathcal{U}\, \psi$ if there exists m such that $n \leq m \leq |w|$, $(w, m) \models \psi$ and, for every k such that $n \leq k < m$, $(w, k) \models \varphi$.

Note that, if $w = w_1 w_1 \cdots w_{|w|}$, $(w, n) \models \varphi$ only depends on the word $w = w_n w_{n+1} \cdots w_{|w|}$.

Example 2.1. Let $w = abbababcba$. Then $(w, 4) \models p_a$ since the fourth letter of w is an a, $(w, 4) \models \circ p_b$ since the fifth letter of w is a b and $(w, 4) \models \diamond(p_c \wedge \circ p_b)$ since cb is a factor of $babcba$.

If φ is a temporal formula, we say that w satisfies φ if $(w, 1) \models \varphi$.

We just have defined "future" temporal formulas but one can define in the same way "past" temporal formulas by reversing time: it suffices to replace "next" by "previous" (symbol $\ominus$), "eventually" by "sometimes" (symbol $\diamondminus$) and "until" by "since" (symbol $\mathcal{S}$). The corresponding semantics are modified as follows.

(3') $(w, n) \models \ominus\varphi$ if $n > 1$ and $(w, n-1)$ satisfies φ.

(4') $(w, n) \models \diamondminus\varphi$ if there exists $m \leq n$ such that $(w, m) \models \varphi$.

(5') $(w, n) \models \varphi \,\mathcal{S}\, \psi$ if there exists $m \leq n$ such that $(w, m) \models \psi$ and for every k such that $m < k \leq n$, $(w, k) \models \varphi$.

The diagram below illustrates the symmetry between the operators "until" and "since".

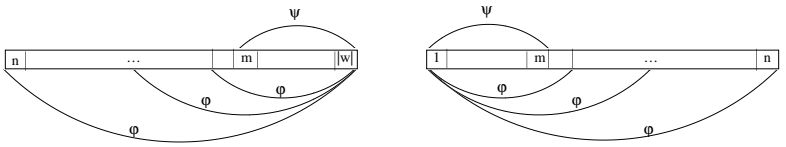


Figure 2.1. A diagram for $(w,n) \models \varphi \,\mathcal{U}\, \psi$ and for $(w,n) \models \varphi \,\mathcal{S}\, \psi$

If φ is a past temporal formula, we say that w satisfies φ if $(w,|w|) \models \varphi$. The language defined by a formula φ is the set $L(\varphi)$ of all words of A^+ that satisfy φ.

3. PTL-definable languages.

In this section, we present a short proof of the following result

Theorem 3.1. *A language of A^+ is PTL-definable if and only if its syntactic semigroup is aperiodic.*

Proof. Since the reverse of an aperiodic semigroup is also aperiodic, it suffices to prove the dual version of the theorem, obtained by using *past* temporal logic. We first prove that every PTL-definable language is star-free (by Schützenberger's theorem, a language is star-free if and only if its syntactic semigroup is aperiodic). This is done by induction on the formation rules. Indeed

(1) $L(p_a) = A^*a$ (for every letter a) is star-free.
(3) $L(\ominus\varphi) = L(\varphi)A$. Thus if $L(\varphi)$ is star-free, so is $L(\ominus\varphi)$.
(4) $L(\diamond\varphi) = L(\varphi)A^*$. Thus if $L(\varphi)$ is star-free, so is $L(\diamond\varphi)$.

We need a similar formula for $\mathcal{S}$, but this is slightly more complicated. Assume that $L(\varphi)$ and $L(\psi)$ are star-free. In particular, there is a semigroup morphism $\eta : A^+ \to S$, where S is an aperiodic semigroup, and a subset P of S such that $L(\varphi) = P\eta^{-1}$. Set, for every $s \in S$, $s^{-1}P = \{t \in S \mid st \in P\}$. Then we have the following lemma, in which $\setminus$ denotes a set difference.

Lemma 3.2. *The following equalities hold*

$$\begin{aligned} L(\varphi \,\mathcal{S}\, \psi) &= \{uv \in A^+ \mid u \in L(\psi), v \in A^* \text{ and for each left factor } v' \neq 1 \text{ of } v, uv' \in L(\varphi)\} \\ &= \bigcup_{s \in S} \Big(s\eta^{-1} \cap L(\psi)\Big)\Big(A^* \setminus (A^+ \setminus (s^{-1}P)\eta^{-1})A^*\Big). \end{aligned}$$

Proof. The first equality is a direct consequence of the definition. Next, if $R \subset A^+$, $(A^+ \setminus R)A^*$ is the set of all words $v \in A^*$ having a left factor $v' \neq 1$ in R. Therefore, taking complements, this is equivalent to saying that $A^* \setminus (A^+ \setminus R)A^*$ is the set of all words $v \in A^*$ such that, for each left factor $v' \neq 1$ of v, $v' \notin R$.

Let $w \in L(\varphi \ \mathcal{S} \ \psi)$. Then, by the first equality, $w = uv$, where $u \in L(\psi)$, $v \in A^*$, and for each left factor $v' \neq 1$ of v, $uv' \in L(\varphi)$. Putting $s = u\eta$, we obtain $u \in s\eta^{-1} \cap L(\psi)$ and $(uv')\eta \in P$, whence $v' \in (s^{-1}P)\eta^{-1}$. Thus,

$$w \in \Big(s\eta^{-1} \cap L(\psi)\Big)\Big(A^* \setminus \big(A^+ \setminus (s^{-1}P)\eta^{-1}\big)A^*\Big)$$

by the remark above.

Conversely, assume that $w = uv$, where, for some $s \in S$, $u \in s\eta^{-1} \cap L(\psi)$ and $v \in A^* \setminus \big(A^+ \setminus (s^{-1}P)\eta^{-1}\big)A^*$. Then $u \in L(\psi)$, and for each left factor $v' \neq 1$ of v, $v' \in (s^{-1}P)\eta^{-1}$. Thus $(uv')\eta = s(v'\eta) \in P$, whence $uv' \in L(\varphi)$. Therefore, by the first equality, $w \in L(\varphi \ \mathcal{S} \ \psi)$. □

Now any language of the form $Q\eta^{-1}$, where $Q \subset S$, is recognized by S, and thus is star-free by Schützenberger's theorem. Therefore, lemma 3.2 shows that $L(\varphi \ \mathcal{S} \ \psi)$ is star-free and this concludes the first part of the proof of theorem 3.1.

We now show that every star-free language is PTL-definable. Let $\mathbf{C}$ be the class of all transformation semigroups X such that *every* language recognized by X is PTL-definable. By Schützenberger's theorem, it suffices to show that each aperiodic semigroup belongs to $\mathbf{C}$. The class $\mathbf{C}$ is certainly closed under division, because if X divides Y, every language recognized by X is also recognized by Y. Next, the trivial semigroup $\{1\}$ belongs to $\mathbf{C}$, since the languages of A^+ recognized by $\{1\}$ are A^+ and the empty set. Now, by theorem 1.5, it remains to show that if $Y = (Q, T) \in \mathbf{C}$, then $U_2 \circ Y \in \mathbf{C}$.

By the wreath-product principle, every language of A^+ recognized by $U_2 \circ Y$ is a finite union of languages of the form $U \cap V\sigma^{-1}$ where $\sigma : A^+ \to B^+ = (Q \times A)^+$ is a certain sequential function, $U \subset A^+$ is recognized by Y and $V \subset B^+$ is recognized by U_2. First, the formulas $L(\neg\varphi) = A^+ \setminus L(\varphi)$ and $L(\varphi \vee \psi) = L(\varphi) \cup L(\psi)$ show that PTL-definable languages are closed under boolean operations. Thus it suffices to show that every language of the form $U \cap V\sigma^{-1}$ above is PTL-definable. Since $Y \in \mathbf{C}$, U is PTL-definable by definition. Furthermore, by lemma 1.3, V is a boolean combination of languages of the form B^*bC^*, where $b \in B$ and $C \subset B$. Since σ^{-1} commutes with boolean operations, it remains to show that languages of the form $(B^*bC^*)\sigma^{-1}$ are PTL-definable. We claim that

$$(B^*bC^*)\sigma^{-1} = (B^*C)\sigma^{-1} \ \mathcal{S} \ (B^*b)\sigma^{-1} \tag{3.1}$$

Indeed, let $u = a_1 \cdots a_n$ be a word of A^+ and let $(a_1 ... a_n)\sigma = b_1 ... b_n$. Then $u\sigma \in B^*bC^*$ if and only if there exists an i such that $b_i = b$ and, for every $j > i$, $b_j \in C$. This is equivalent to saying that $(a_1 \cdots a_i)\sigma \in B^*b$ and for every $j > i$, $(a_1 \cdots a_j)\sigma \in B^*C$, and this proves (3.1). Now

$$(B^*C)\sigma^{-1} = \bigcup_{b \in C} (B^*b)\sigma^{-1}$$

and therefore it suffices to show that languages of the form $(B^*b)\sigma^{-1}$ are PTL-definable. We take again the notations used in the definition of σ (cf. proposition 1.8). Set $b = (q, a)$ (recall that $B = Q \times A$). Then we have

$$(a_1 \cdots a_n)\sigma = (q_0, a_1)(q_0 \cdot (a_1\eta\pi), a_2) \cdots (q_0 \cdot (a_1 \cdots a_{n-1})\eta\pi, a_n).$$

It follows that $(a_1 ... a_n)\sigma \in B^*b$ if and only if $q_0 \cdot (a_1 \cdots a_{n-1})\eta\pi = q$ and $a_n = a$. Therefore $(B^*b)\sigma^{-1} = La$, where $L = \{u \in A^+ \mid q_0 \cdot (u\eta\pi) = q\}$. But L is recognized by Y and since $Y \in \mathbf{C}$, is PTL-definable. Now, since $L(\varphi)a = L(\ominus\varphi \wedge p_a)$, $La = (B^*b)\sigma^{-1}$ is PTL-definable and this concludes the proof. □

4. Restricted temporal logic.

If we omit the "until" operator, we obtain a restricted temporal logic (RTL) that was considered in [5,6]. Here is a first description of the languages definable in this logic. The subtle distinction between conditions (2) and (3) will be used in the proof of the main theorem below.

Proposition 4.1. *Let L be a language of A^+. The following conditions are equivalent:*

(1) L is RTL-definable,

(2) L belongs to the smallest boolean algebra of languages containing the languages aA^* and closed under the operations $L \to A^*L$ and $L \to aL$ for every $a \in A$,

(3) L belongs to the smallest boolean algebra of languages containing the languages aA^* and closed under the operations $L \to A^*aL$ and $L \to aL$ for every $a \in A$.

Proof. Let $\mathcal{C}$ (respectively $\mathcal{C}'$) be the smallest boolean algebra of languages closed under the operations $L \to A^*L$ (respectively $L \to A^*aL$) and $L \to aL$ for every letter $a \in A$. In particular the languages $\emptyset$ and A^+ belong to $\mathcal{C}$ and $\mathcal{C}'$ by definition. We first prove that $\mathcal{C} = \mathcal{C}'$. The inclusion $\mathcal{C}' \subset \mathcal{C}$ follows directly from the formula $A^*aL = A^*(aL)$. The opposite inclusion follows from the formula

$$A^*L = L \cup \bigcup_{a \in A} A^*aL.$$

Thus (2) and (3) are equivalent.

(1) implies (2). We show by induction on the formation rules that $L(\varphi) \in \mathcal{C}$ for every RTL-formula φ. First, if $\varphi = p_a$, then

$$L(p_a) = aA^* \in \mathcal{C}.$$

If φ and ψ are formulas such that $L(\varphi)$ and $L(\psi)$ belong to $\mathcal{C}$, then

$$\begin{aligned} L(\varphi \vee \psi) &= L(\varphi) \cup L(\psi) \in \mathcal{C}, \\ L(\varphi \wedge \psi) &= L(\varphi) \cap L(\psi) \in \mathcal{C}, \\ L(\neg\varphi) &= A^+ \setminus L(\varphi) \in \mathcal{C}, \\ L(\circ\varphi) &= AL(\varphi) = \bigcup_{a \in A} aL(\varphi) \in \mathcal{C}, \\ L(\diamond\varphi) &= A^*L(\varphi) \in \mathcal{C}. \end{aligned}$$

(2) implies (1). Let $\mathcal{F}$ be the set of RTL-definable languages. Then $\mathcal{F}$ contains $aA^* = L(p_a)$, for every $a \in A$. The formulas $L(\varphi) \cup L(\psi) = L(\varphi \vee \psi)$ and $A^+ \setminus L(\varphi) = L(\neg\varphi)$ show that $\mathcal{F}$ is a boolean algebra and the formula $A^*L(\varphi) = L(\diamond\varphi)$ shows that $\mathcal{F}$ is closed under the operation $L \to A^*L$. Finally, the formula $aL(\varphi) = L(p_a \wedge \circ\varphi)$ shows that $\mathcal{F}$ is closed under the operation $L \to aL$, for every letter $a \in A$. Therefore $\mathcal{F}$ contains $\mathcal{C}$. □

We can now state our main result.

Theorem 4.2. *Let L be a language of A^+. The following conditions are equivalent:*
(1) L is RTL-definable,
(2) the syntactic semigroup of L is locally $\mathcal{L}$-trivial.

Proof. As for theorem 3.1, we prove the dual version of the theorem, which states that L is definable in past restricted temporal logic if and only if its syntactic semigroup is $\mathcal{R}$-trivial. Consider the smallest boolean algebra $\mathcal{B}$ containing the languages A^*a and closed under the operations $L \to LaA^*$ and $L \to La$ for every $a \in A$. By proposition 4.1 and duality, it suffices now to prove the following statement

"A language belongs to $\mathcal{B}$ if and only if its syntactic semigroup belongs to **LR**".

First, $S(A^*a) = B(1,2) \in \mathbf{LR}$. Now, by proposition 1.7, $S(LaA^*)$ divides $U_1 \circ S(L)$, and $S(La)$ divides $B(1,2) \circ S(L)$. It follows by proposition 1.10, that if $S(L) \in \mathbf{LR}$, then $S(LaA^*) \in \mathbf{LR}$ and $S(La) \in \mathbf{LR}$. Therefore, if $L \in \mathcal{B}$, then $S(L) \in \mathbf{LR}$.

In the other direction, the proof mimics the proof of theorem 3.1. Let $\mathbf{C}$ be the class of all transformation semigroups X such that *every* language recognized by X belongs to $\mathcal{B}$. The class $\mathbf{C}$ contains the trivial semigroup and is closed under division. Therefore, to show that $\mathbf{C}$ contains **LR**, it suffices, by proposition 1.10, to verify that if $Y \in \mathbf{C}$, then $U_1 \circ Y \in \mathbf{C}$ and $\mathbf{2} \circ Y \in \mathbf{C}$.

By the wreath-product principle, every language of A^+ recognized by $U_1 \circ Y$ (respectively $\mathbf{2} \circ Y$) is a finite union of languages of the form $U \cap V\sigma^{-1}$ where $\sigma : A^+ \to B^+ = (Q \times A)^+$ is a certain sequential function, $U \subset A^+$ is recognized by Y and $V \subset B^+$ is recognized by U_1 (respectively **2**). Since $Y \in \mathbf{C}$, U belongs to $\mathcal{B}$ by definition. Furthermore, by lemma 1.3, V is a boolean combination of languages of the form B^*bB^*, (respectively B^*b) where $b \in B$. Since σ^{-1} commutes with boolean operations, it remains to show that the languages of the form $(B^*bB^*)\sigma^{-1}$ (respectively $(B^*b)\sigma^{-1}$) belong to $\mathcal{B}$. We take again the notations used in the definition of σ (cf. proposition 1.8). Set $b = (q, a)$ (recall that $B = Q \times A$). Then we have

$$(a_1 \cdots a_n)\sigma = (q_0, a_1)(q_0 \cdot (a_1\eta\pi), a_2) \cdots (q_0 \cdot (a_1 \cdots a_{n-1})\eta\pi, a_n).$$

First assume $q \neq q_0$. Then $(a_1 ... a_n)\sigma \in B^*bB^*$ if and only if there exists an index i such that $q_0 \cdot (a_1 \cdots a_{i-1})\eta\pi = q$ and $a_i = a$. Therefore $(B^*bB^*)\sigma^{-1} = LaA^*$, where

$$L = \{u \in A^+ \mid q_0 \cdot (u\eta\pi) = q\}.$$

If $q = q_0$, then $(B^*bB^*)\sigma^{-1} = LaA^* \cup aA^*$. But L is recognized by Y and since $Y \in \mathbf{C}$, L belongs to $\mathcal{B}$. Furthermore, aA^* also belongs to $\mathcal{B}$, since

$$\{a\} = A^*a \setminus \big((A^*a)aA^* \cup (A^*a)bA^* \cup (A^*b)aA^* \cup (A^*b)bA^*\big),$$

and

$$aA^* = \{a\} \cup \{a\}aA^* \cup \{a\}bA^*$$

It follows that $(B^*bB^*)\sigma^{-1}$ belongs to $\mathcal{B}$. Similarly, $(a_1 \cdots a_n)\sigma \in B^*b$ if and only if $q_0 \cdot (a_1 \cdots a_{n-1})\eta\pi = q$ and $a_n = a$. Therefore $(B^*b)\sigma^{-1} = La$ or $La \cup \{a\}$ (if $q = q_0$) and $(B^*b)\sigma^{-1}$ also belongs to $\mathcal{B}$. □

Corollary 4.3. *Given a rational language L, one can effectively decide whether it is RTL-definable.*

Proof. The language L can be given either by a rational expression or by a finite automaton. In both cases, there are well-known algorithms to compute its minimal automaton $\mathcal{A}(L)$, and then its syntactic semigroup $S(L)$, which is also the transformation semigroup of $\mathcal{A}(L)$. Then it suffices, by proposition 1.9 to verify that $S(L)$ satisfies the identity $x^\omega v(x^\omega u x^\omega v)^\omega = (x^\omega u x^\omega v)^\omega$. □

Say that two PTL-formulas φ and ψ are *equivalent* if $L(\varphi) = L(\psi)$, that is, if they agree when interpreted on finite words.

Corollary 4.4. *Given a PTL-formula, one can effectively decide whether it is equivalent to some RTL-formula.*

We conclude this section by three examples.

Example 4.1. Let $A = \{a, b\}$ and let $L = (ab)^+$. Then the minimal automaton of L is represented in the diagram below.

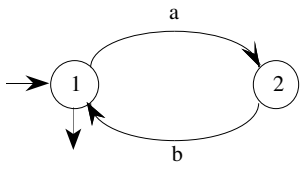


Figure 4.1.

The syntactic semigroup of L is the semigroup S with zero presented by the relations

$$a^2 = b^2 = 0, \quad aba = a, \quad bab = b.$$

Thus $S = \{a, b, ab, ba, 0\}$. There are three idempotents ab, ba, and 0. The corresponding "local" semigroups are

$$abSab = \{ab, 0\}, \quad baSba = \{ba, 0\} \quad \text{and} \quad 0S0 = \{0\},$$

all of which are $\mathcal{L}$-trivial. Therefore L is expressible in restricted temporal logic. Indeed, we have $L = L(\varphi)$, where

$$\varphi = p_a \wedge \Diamond(p_b \wedge \neg \circ p_a \wedge \neg \circ p_b) \wedge \neg \Diamond (p_a \wedge \circ p_a) \wedge \neg \Diamond (p_b \wedge \circ p_b).$$

Example 4.2. Let $A = \{a, b, c\}$ and let $L = A^* a \{a, c\}^*$. Then the minimal automaton of L is represented in the diagram below.

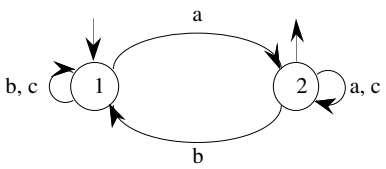


Figure 4.2.

The syntactic semigroup of L is U_2, which is locally $\mathcal{L}$-trivial. Therefore L is expressible in restricted temporal logic. Indeed, we have $L = L(\varphi)$, where

$$\varphi = \Diamond(p_a \wedge \circ\neg \Diamond p_b).$$

Example 4.3. Let $A = \{a, b, c\}$ and let $L = a^* b \{a, b, c\}^*$. Then the minimal automaton of L is represented in the diagram below.

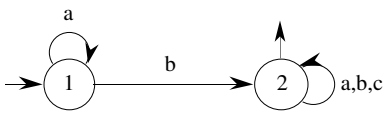


Figure 4.3.

The syntactic semigroup of L is the monoid S presented by the relations

$$a = 1, \quad bb = bc = b, \quad cb = cc = c.$$

This is the reverse of U_2, and it is aperiodic, but not locally $\mathcal{L}$-trivial. Therefore, any formula ψ such that $L = L(\psi)$ uses the "until" operator. In fact, $L = L(\varphi)$, where $\varphi = p_a \,\mathcal{U}\, p_b$.

5. Automata, varieties and forbidden configurations.

In the two previous sections, we have seen how to characterize the formal languages associated with a formula of propositional temporal logic (section 3) and of restricted temporal logic (section 4). Both characterizations are in terms of the syntactic semigroup of the formal language. We shall see here how this characterization can be expressed in terms of automata. In the case of restricted temporal logic, this has the advantage of providing a polynomial algorithm to check whether the language defined by a given deterministic automaton is RTL-definable. This is of interest since, on the contrary, the corresponding problem for PTL logic is the complement of an NP-hard problem [19] and is PSPACE-complete [1]. Thus, unless $P = NP$, checking whether the language defined by a given automaton is PTL-definable cannot be solved in polynomial time.

We begin with the characterization of automata associated with $\mathcal{R}$-trivial semigroups. We shall then treat the case of locally $\mathcal{R}$-trivial semigroups. This corresponds, as we have seen, to formulas of past temporal logic. We shall finally come to $\mathcal{L}$-trivial and locally $\mathcal{L}$-trivial semigroups, which correspond to RTL-formulas. We shall see how these characterizations lead to polynomial algorithms.

Before to give the details of our algorithms, let us fix some convenient notations. 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 $\mathcal{A}^k$. For instance, if $\mathcal{A}$ is the automaton represented below

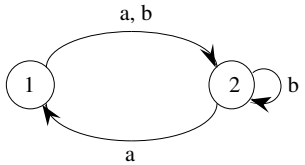


Figure 5.1.

then $\mathcal{A}^2$ is the automaton

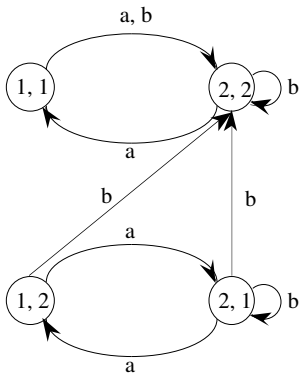


Figure 5.2.

and $G_2(\mathcal{A})$ is the graph

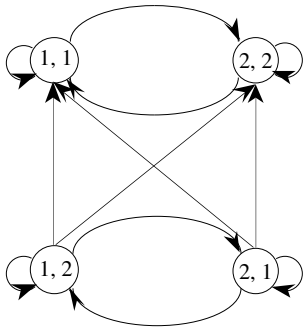
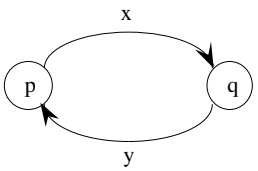


Figure 5.3.

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 $\mathcal{A}$. Two words $x, y \in A^*$ which have the same action on Q are said to be equivalent in $\mathcal{A}$ (notation $x \equiv y$). The following result is already in [15], p. 118.

Theorem 5.1. *The semigroup of a deterministic automaton $\mathcal{A}$ is $\mathcal{R}$-trivial if and only if there exist no configurations of $\mathcal{A}$ of the form*



with $p \neq q$.

Proof. Suppose first that $S(\mathcal{A})$ is $\mathcal{R}$-trivial and consider a configuration as above. Let ω be the exponent of $S(\mathcal{A})$. Then we have, for every $x, y \in A^+$

$$(xy)^\omega \equiv (xy)^\omega x$$

and therefore

$$p = p \cdot (xy)^\omega = p = p \cdot (xy)^\omega x = q$$

whence $p = q$. Conversely, if $\mathcal{A} = (Q, A, \cdot)$ contains no forbidden configuration, let us verify that, for every $u, v \in A^+$, $(uv)^\omega \equiv (uv)^\omega u$. Let $r \in Q$ and let $p = r \cdot (uv)^\omega$. Since $(uv)^\omega$ is idempotent, we have $p \cdot (uv)^\omega = p$. Set $x = u$, $y = (vu)^{\omega-1}v$ and $q = p \cdot x$. Then $q \cdot y = p \cdot xy = p \cdot (uv)^\omega = p$. Therefore $\mathcal{A}$ contains the configuration

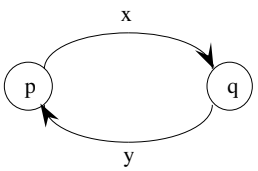


Figure 5.4.

and thus $p = q$. Therefore $p = r \cdot (uv)^\omega = r \cdot (uv)^\omega u$ and thus $(uv)^\omega \equiv (uv)^\omega u$. □

The transposition of the previous characterization to the case of locally $\mathcal{R}$-trivial semigroups follows a general scheme. Let $\mathbf{V}$ be a variety of semigroups and assume that the deterministic automata whose semigroups belong to $\mathbf{V}$ can be described by a set $\mathcal{C}$ of forbidden configurations. Then the deterministic automata whose semigroups belong to the variety $\mathbf{LV}$ of all semigroups which are locally in $\mathbf{V}$ can be described by the set $\mathcal{C}'$ of forbidden configurations obtained as follows. For each configuration $C \in \mathcal{C}$, we add to each vertex a loop labeled by a new symbol, the same for all vertices. Then the semigroup of a deterministic automaton $\mathcal{A}$ belongs to $\mathbf{LV}$ if and only if that $\mathcal{A}$ contains no configuration of $\mathcal{C}'$.

In particular, we have the following result.

Theorem 5.2. *The semigroup of a deterministic automaton $\mathcal{A}$ is locally $\mathcal{R}$-trivial if and only if there exist no configurations of $\mathcal{A}$ of the form*

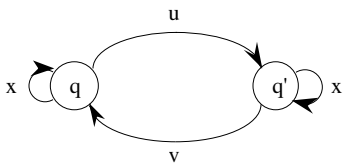


Figure 5.5.

with $q \neq q'$.

Proof. By proposition 1.9, a semigroup is locally $\mathcal{R}$-trivial if and only if it satisfies the identity

$$(ux^\omega vx^\omega)^\omega ux^\omega = (ux^\omega vx^\omega)^\omega \tag{5.1}$$

Suppose that $S(\mathcal{A})$ is locally $\mathcal{R}$-trivial and that $\mathcal{A}$ contains a configuration of the form represented in 5.5. Then by (5.1),

$$q = q \cdot (ux^\omega vx^\omega)^\omega = q \cdot (ux^\omega vx^\omega)^\omega ux^\omega = q'$$

Conversely, suppose that $\mathcal{A}$ satisfies the condition of the theorem, and let u, v, x be arbitrary words of A^+. Set $u' = ux^\omega$, $v' = vx^\omega$ and $x' = x^\omega$. Let q be a state, and set $q_1 = q \cdot (ux^\omega vx^\omega)^\omega$ and $q_2 = q_1 \cdot ux^\omega$. Then a short computation shows that $\mathcal{A}$ contains the following configuration:

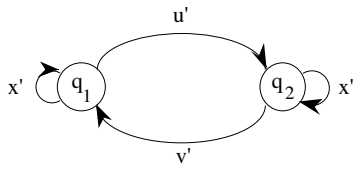


Figure 5.6.

and thus $q_1 = q_2$. It follows that $q \cdot (ux^\omega vx^\omega)^\omega = q \cdot (ux^\omega vx^\omega)^\omega ux^\omega$ for any state q, and thus S satisfies the identity (5.1). Thus $S(\mathcal{A})$ is locally $\mathcal{R}$-trivial. □

The previous result yields to a polynomial time algorithm to check whether the semigroup of an n-state deterministic automaton $\mathcal{A}$ is locally $\mathcal{R}$-trivial or not. Indeed, one first observe that given two states q and q', there is a word $w \in A^+$ such that $q \cdot w = q$ and $q' \cdot w = q'$ if and only if $\big((q,q'),(q,q')\big)$ is an edge in the directed graph $G_2(\mathcal{A})$. Therefore, one can check whether $\mathcal{A}$ contains a configuration of the form 5.5 with $q \neq q'$ by computing G_1 and G_2 and by verifying there are no pairs $\{q, q'\}$ of states such that

(a) (q,q') and (q',q) are edges in $G_1(\mathcal{A})$, and

(b) $\big((q,q'),(q,q')\big)$ is an edge of $G_2(\mathcal{A})$.

Since G_1 (resp. G_2) has n (n^2) vertices, this gives a polynomial algorithm.

This is in fact a general property of varieties defined by forbidden configurations. Let indeed $\mathbf{V}$ be a variety of semigroups and assume that the deterministic automata whose semigroups belong to $\mathbf{V}$ can be described by a finite set $\mathcal{C}$ of forbidden configurations. Then there is a polynomial algorithm to check whether a given n-state deterministic automaton $\mathcal{A}$ belongs to $\mathbf{V}$. For this we have to check whether or not some configuration C of $\mathcal{C}$ is present in $\mathcal{A}$. The number of possible assignements of states to the vertices of C is polynomial in n. And for each assignement, the existence of a given set of k edges with the same label is solved by reduction to an accessibility problem in the graph $G_k(\mathcal{A})$. The overall algorithm is polynomial.

In particular, we have the following result.

Corollary 5.3. *There is a polynomial time algorithm for testing whether the reverse of the language accepted by an n-state deterministic automaton is RTL-definable.*

We illustrate this method on the following example.

Example 5.1. Let $\mathcal{A}$ be the automaton given on figure 5.7 and already considered in example 4.2

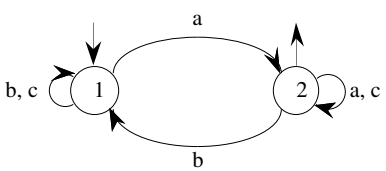


Figure 5.7.

To check whether $S(\mathcal{A})$ is locally $\mathcal{R}$-trivial, we construct the graph $G_2(\mathcal{A})$. It is represented in Figure 5.8

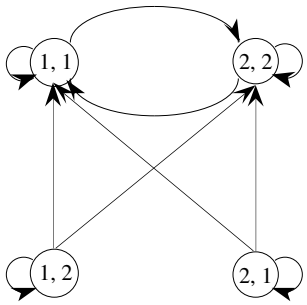


Figure 5.8.

Now this graph contains a cycle of length 1 around (1,2) and 1 and 2 are in the same strongly connected component of $G_1(\mathcal{A})$. This indicates the presence of a forbidden configuration. It is indeed obtained for instance with the labels given in figure 5.9

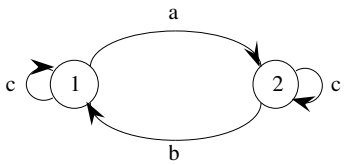
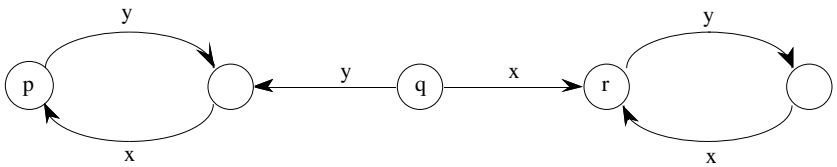


Figure 5.9.

It follows that $\mathcal{A}$ is not $\mathcal{R}$-trivial and $L(\mathcal{A})$ is not expressible in reverse restricted temporal logic.

We now consider the case of $\mathcal{L}$-trivial semigroups.

Proposition 5.4. *The semigroup of a deterministic automaton $\mathcal{A}$ is $\mathcal{L}$-trivial if and only if the configuration*



with $p \neq r$ is not present in $\mathcal{A}$.

Proof. Let us first suppose that $S(\mathcal{A})$ is $\mathcal{L}$-trivial. We consider a configuration as above. Since $(yx)^\omega \equiv x(yx)^\omega$, we have

$$r = q \cdot x(yx)^\omega = q \cdot (yx)^\omega = p$$

whence $p = r$. Conversely, suppose that the above configuration is not present in $\mathcal{A}$. Let $x, y \in A^+$ and let $q \in Q$ be arbitrary. Let $r = q \cdot x(yx)^\omega$ and $p = q \cdot (yx)^\omega$. Then $p = r$ by the hypothesis and therefore $x(yx)^\omega \equiv (yx)^\omega$. Thus $S(\mathcal{A})$ is $\mathcal{L}$-trivial. □

Note that the characterization of proposition 5.4, contrary to that of proposition 5.1 requires the hypothesis that the automaton is complete. There is in fact no possibility of characterization by forbidden configurations of $\mathcal{L}$-trivial semigroups given by a deterministic automaton if it is not complete. Indeed the automaton of figure 5.10 (i) is a subgraph of the labeled graph of the automaton of figure 5.10 (ii). The semigroup of the first one is not $\mathcal{L}$-trivial whereas the second one is.

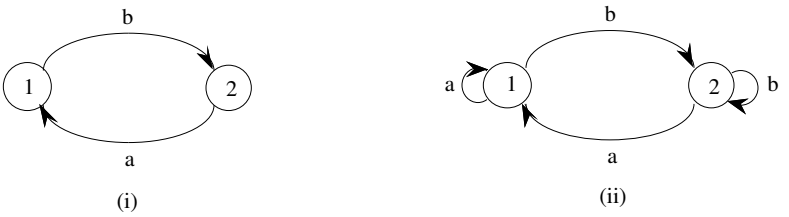
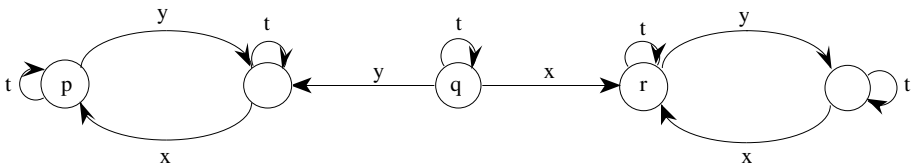


Figure 5.10.

We finally give the announced characterization of locally $\mathcal{L}$-trivial semigroups. It is a corollary of proposition 5.4.

Proposition 5.5. *The semigroup of a deterministic automaton $\mathcal{A}$ is locally $\mathcal{L}$-trivial if and only if the configuration*



with $p \neq r$ is not present in $\mathcal{A}$.

Together with theorem 4.2, we obtain.

Corollary 5.6. *There is a polynomial time algorithm for testing whether the language accepted by an n-state deterministic automaton is RTL-definable.*

This does not give, however, a polynomial algorithm to check whether a given PTL-formula is equivalent with a RTL formula. We presently do not know any reasonable bound on the complexity of this problem.

6. Conclusion.

We have given an effective characterization of the languages definable in linear propositional temporal logic and in restricted temporal logic. It would be interesting to obtain similar characterizations when the temporal logic is interpreted on infinite words. This will be the subject of a future paper. Another interesting question is to consider the temporal logic whose only operator is "eventually". Sistla and Zuck [18] have given a description of the set of infinite words definable in this logic, but this description doesn't seem to be effective.

Acknowledgements.

We would like to thank H. Straubing for some useful comments on an earlier version of this work.

7. Bibliography

[1] Sang Cho and Dung T. Huynh, Finite-automaton aperiodicity is PSPACE-complete, *Theoretical Computer Science* **88** (1991), 99–116.

[2] R.S. Cohen and J.A. Brzozowski, On star-free events, *Proc. Hawaii Internat. Conf. Syst. Sci.*, Honolulu, (1968), 1–4.

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

[4] E.H. Emerson, J.Y. Halpern, "Sometimes" and "not never" revisited: On Branching vs. Linear Time, *J. Assoc. Comput. Math.* **33**, (1986) 151–178.

[5] D. Gabbay, A. Pnueli, S. Shelah, J. Stavi, On the temporal analysis of fairness, Proc. *12th ACM Symp. on Principles of Programming Languages*, Las Vegas, (1980), 163–173

[6] J.A. Kamp, *Tense logic and the theory of linear order*, Ph. D. Thesis, University of California, Los Angeles, (1968).

[7] O. Katai, Completeness and the expressive power of nexttime temporal logical system by semantic tableau method, INRIA report **109**, (1981).

[8] G. Lallement, *Semigroups and combinatorial applications*, Wiley, New-York, (1979).

[9] A.R. Meyer, A note on star-free events, *J.ACM* **16**, (1969), 220–225.

[10] R. McNaughton and S. Papert, *Counter-free automata*, MIT Press, Cambridge, Mass, (1971).

[11] M. Parigot, Automates, réseaux, formules. Actes des Journées "Informatique et Mathématiques", Luminy (1984), 74–89

[12] R. Peikert, ω-regular languages and propositional temporal logic, preprint.

[13] D. Perrin and J.E. Pin, First order logic and star-free sets, *J. Comput. System Sci.* **32**, 1986, 393–406.

[14] D. Perrin, *Finite automata*, in Handbook of Theoret. Comput. Sci., Vol. B, J. van Leuwen ed., North Holland, (1990) 2–57.

[15] J.E. Pin, *Varieties of formal languages*, North Oxford Academic, London and Plenum, New-York, 1986

[16] A. Pnueli, The temporal logic of programs, *Proc. 18th FOCS*, Providence, RI, (1977), 46–57.

[17] M.P. Schützenberger, On finite monoids having only trivial subgroups, *Inform. and Control* **48**, (1965), 190–194.

[18] A.P. Sistla and L.D. Zuck, On the eventually operator in temporal logic, *in Proceedings Second IEEE Symposium on Logic in Computer Science* (1987) 153–166.

[19] J. Stern, Complexity of some problems from the theory of automata, *Information and Control* **66** (1985) 163–176.

[20] P.Stiffler, Extension of the Fondamental Theorem of Finite Semigroups, *Advances in Mathematics* **11**, (1973), 159–209.

[21] H. Straubing, Varieties of recognizable sets whose syntactic monoids contain solvable groups, *Ph. D. Thesis*, University of California, Berkeley, CA, (1978).

[22] H. Straubing, Finite semigroup varieties of the form $V * D$, *J. Pure Applied Algebra* **36** (1985) 53–94.

[23] J. van Leeuwen, Graph algorithms, Chap. 10 *in* Handbook of Theoretical Computer Science, Edited by J. van Leeuwen, Elsevier Science Publishers B.V., (1990) Vol. A, 525–631.

[24] M.Y. Vardi and P. Wolper, Applications of temporal logic: an automata-theoretic perspective, preprint, (1985).