Semigroups and automata on infinite words

Dominique Perrin*and Jean-Éric Pin†

Published in 1995

1 Introduction

This paper is an introduction to the algebraic theory of infinite words. Infinite words are widely used in computer science, in particular to model the behaviour of programs or circuits. From a mathematical point of view, they have a rich structure, at the cross-roads of logic, topology and algebra. This paper emphasizes the combinatorial and algebraic aspects of this theory but the interested reader is referred to the survey articles [34, 44] or to the report [30] for more information on the other aspects. In particular, the important topic of the complexity of the algorithms on infinite words is not treated in this paper.

The paper is written with the perspective of generalizing the results on recognizable sets of finite words to infinite words. This does not exactly follow the historical development of the theory, but it gives a good idea of the type of problems that occur in this field. Some of these problems are still open, or have been solved quite recently so that the definitions and results presented below may not be as yet finalized.

The first result to be generalized is the equivalence between finite automata, finite deterministic automata and rational expressions. If one adds infinite iteration ("omega" operation) to the standard rational operations, union, product and star, one gets a natural definition of the ω-rational sets of infinite words that extends the definition of rational sets of finite words. Büchi [5] was the first to propose a definition of finite automata acting on infinite words. This definition suffices to extend Kleene's theorem to infinite words: the sets of infinite words recognized by finite Büchi automata are exactly the ω-rational sets. This result is now known as Büchi's theorem.

However, Büchi's definition is not totally satisfying since deterministic Büchi automata are not equivalent to non deterministic ones. The con-

*Institut Gaspard Monge, Université de Marne la Vallée, 93166 Noisy-le-Grand, FRANCE.

†From 1st Sept 1993, LITP, Université Paris VI, Tour 55-65, 4 Place Jussieu, 75252 Paris Cedex 05, FRANCE. **E-mail**: `pin@litp.ibp.fr`

nection between deterministic and non deterministic Büchi automata was enlightened by a deep theorem of McNaughton: a set of infinite words is recognized by a non deterministic Büchi automaton if and only if it is a finite boolean combination of sets recognized by deterministic Büchi automata. This result was prepared by a suitable definition for automata on infinite words given by Muller [22]. These automata have the same power as the Büchi automata, but this time, non deterministic automata are equivalent to deterministic ones.

It is a well known fact that finite semigroups can be viewed as a two-sided algebraic counterpart of finite automata that recognize finite words. Several attempts have been made to find an algebraic counterpart of finite automata that recognize infinite words. Since the notion of infinite word is asymmetrical, finite semigroups are not suitable any more. They can be replaced by ω-semigroups, which are, roughly speaking, semigroups equipped with an infinite product. The basic definitions are quite promising, but a technical difficulty arises almost immediately. Indeed, in order to design algorithms on finite ω-semigroups one needs a finite representation for them, but the definition of the infinite product apparently requires an infinite table, even for a 2-element ω-semigroup. However, a Ramsey-type argument shows that the structure of a finite ω-semigroup is totally determined by three operations of finite signature. That is, finite ω-semigroupsare equivalent to certain finite algebras of finite signature, the Wilke algebras. Now, the definitions of a recognizable set, of a syntactic congruence, etc., become natural and most results valid for finite words can be adapted to infinite words. Carrying on the work of Arnold [1], Pécuchet [24, 23] and the first author [25, 26, 27], Wilke [45, 46] has pushed the analogy with the theory for finite words sufficiently far to obtain a counterpart of Eilenberg's variety theorem for finite or infinite words. This is approximatively the point reached by the algebraic approach today, although current research may already have passed beyond. In any case, this is the place where our article finishes.

The paper is organized as follows: the basic definitions on words and ω-rational sets are given in sections 2 and 3. Büchi automata are defined in section 4, deterministic Büchi automata in section 5 and Muller automata in section 6. The equivalence between finite ω-semigroups and finite Wilke algebras is established in section 7 and their connections with automata are presented in sections 8, 9 and 10. Syntactic ω-semigroups are introduced in section 11 and the variety theorem and its consequences are discussed in section 12. Section 13 presents the conclusion of the article.

2 Words

Let A be a finite set called an *alphabet*, whose elements are *letters*. A *finite word* is a finite sequence of letters, that is, a function u from a finite set of the form $\{0, 1, 2, \ldots, n\}$ into A. If one puts $u(i) = a_i$ for $0 \leqslant i \leqslant n$, the word u is usually denoted by $a_0 a_1 \cdots a_n$, and the integer $|u| = n + 1$ is the *length* of u. The unique word of length 0 is the empty word, denoted by 1. An *infinite word* is a function u from $\mathbb{N}$ into A, usually denoted by $a_0 a_1 a_2 \cdots$, where $u(i) = a_i$ for all $i \in \mathbb{N}$. A *word* is either a finite word or an infinite word.

Intuitively, the *concatenation* or *product* of two words u and v is the word uv obtained by writing u followed by v. More precisely, if u is finite and v is finite or infinite, then uv is the word defined by

$$(uv)(i) = \begin{cases} u(i) & \text{if } i < |u| \\ v(i - |u|) & \text{if } i \geqslant |u| \end{cases}$$

We denote respectively by A^*, A^+, $A^{\mathbb{N}}$ the set of all finite words, finite non-empty words, and infinite words, respectively. We also denote by $A^\infty = A^+ \cup A^{\mathbb{N}}$ the set of all non empty words. A word x is a factor of a word w if there exist two words u and v (possibly empty) such that $w = uxv$.

3 Rational sets

The *rational operations* are the four operations union, product, plus and star, defined on the set of subsets of A^* as follows

(1) Union: $L_1 \cup L_2 = \{u \mid u \in L_1 \text{ or } u \in L_2\}$

(2) Product: $L_1 L_2 = \{u_1 u_2 \mid u_1 \in L_1 \text{ and } u_2 \in L_2\}$

(3) Plus: $L^+ = \{u_1 \cdots u_n \mid n > 0 \text{ and } u_1, \ldots, u_n \in L\}$

(4) Star: $L^* = \{u_1 \cdots u_n \mid n \geqslant 0 \text{ and } u_1, \ldots, u_n \in L\}$

Thus we have the relations

$$L^+ = LL^* = L^*L \quad \text{and} \quad L^* = L^+ \cup \{1\}$$

The set of rational subsets of A^* is the smallest set of subsets of A^* containing the finite sets and closed under finite union, product and star. For instance, $(a \cup ab)^*ab \cup (ba^*b)^*$ denotes a rational set.

Similarly, the set of rational subsets of A^+ is the smallest set of subsets of A^+ containing the finite sets and closed under finite union, product and plus. It is not difficult to verify that the rational subsets of A^+ are exactly the rational subsets of A^* that do not contain the empty word.

It is possible to generalize the concept of rational sets to infinite words as follows. First, the product can be extended, by setting, for $X \subset A^*$ and

$Y \subset A^{\mathbb{N}}$,

$$XY = \{xy \mid x \in X \text{ and } y \in Y\}.$$

Next, we define an infinite iteration ω by setting, for every subset X of A^+

$$X^\omega = \{x_0 x_1 \cdots \mid \text{ for all } i \geqslant 0, x_i \in X\}$$

Equivalently, X^ω is the set of infinite words obtained by concatenating an infinite sequence of words of X. In particular, if $u = a_0 a_1 \cdots a_n$, we set

$$u^\omega = a_0 a_1 \cdots a_n a_0 a_1 \cdots a_n a_0 a_1 \cdots a_n a_0 a_1 \cdots$$

The set $\mathcal{R}at(A^\infty)$ of *ω-rational* subsets of A^∞ is the smallest set $\mathcal{R}$ of subsets of A^∞ such that

(a) $\emptyset \in \mathcal{R}$ and for all $a \in A$, $\{a\} \in \mathcal{R}$,

(b) $\mathcal{R}$ is closed under finite union,

(c) For every subset X of A^+ and for every subset Y of A^∞, $X \in \mathcal{R}$ and $Y \in \mathcal{R}$ imply $XY \in \mathcal{R}$,

(d) For every subset X of A^+, $X \in \mathcal{R}$ implies $X^+ \in \mathcal{R}$ and $X^\omega \in \mathcal{R}$.

In other words, the set of ω-rational subsets of A^∞ is the smallest set of subsets of A^∞ containing the finite sets of A^+ and closed under finite union, finite product, plus and omega. The ω-rational sets which are contained in $A^{\mathbb{N}}$ are called, by abuse of language, the ω-rational subsets of $A^{\mathbb{N}}$. There is a very simple characterization of these sets, which is often used as a definition.

Proposition 3.1 *A subset of $A^{\mathbb{N}}$ is ω-rational if and only if it is a finite union of subsets of the form XY^ω where X and Y are non-empty rational subsets of A^+.*

Example 3.1 The set of infinite words on the alphabet $\{a, b\}$ having only a finite number of b's is given by the expression $\{a, b\}^* a^\omega$.

4 Automata

A finite (non deterministic) automaton is a triple $\mathcal{A} = (Q, A, E)$ where Q is a finite set (the set of *states*), A is an alphabet, and E is a subset of $Q \times A \times Q$, called the set of *transitions*. Two transitions (p, a, q) and (p', a', q') are *consecutive* if $q = p'$. A *path* in $\mathcal{A}$ is a finite sequence of consecutive transitions

$$e_0 = (q_0, a_0, q_1),\ \ e_1 = (q_1, a_1, q_2),\ \ \ldots\ ,\ \ e_{n-1} = (q_{n-1}, a_{n-1}, q_n)$$

also denoted

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

The state q_0 is the *origin* of the path, the state q_{n+1} is its *end*, and the word $x = a_0 a_1 \cdots a_n$ is its *label*.

An *infinite path* in $\mathcal{A}$ is a sequence p of consecutive transitions indexed by $\mathbb{N}$.

$$e_0 = (q_0, a_0, q_1), \ \ e_1 = (q_1, a_1, q_2), \ \ \ldots$$

also denoted

$$q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \ \cdots$$

The state q_0 is the *origin* of the infinite path and the infinite word $a_0 a_1 \cdots$ is its *label*. A state q occurs infinitely often in p if $q_n = q$ for infinitely many n.

Example 4.1 Let $\mathcal{A} = (Q, A, E)$ be the automaton represented in Figure 4.1. Then $Q = \{1, 2\}$, $A = \{a, b\}$, $E = \{(1, a, 1), (2, b, 1), (1, a, 2), (2, b, 2)\}$ and

$$(1, a, 2)(2, b, 2)(2, b, 1)(1, a, 2)(2, b, 2)(2, b, 1)(1, a, 2)(2, b, 2)(2, b, 1)(1, a, 2)\cdots$$

is an infinite path of $\mathcal{A}$.

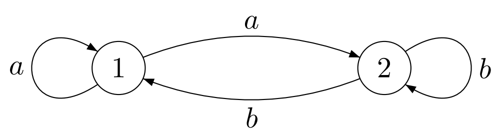

Figure 4.1: A finite automaton.

A finite *Büchi automaton* is a quintuple $\mathcal{A} = (Q, A, E, I, F)$ where

(1) (Q, A, E) is a finite automaton,

(2) I and F are subsets of Q, called the set of *initial* and *final* states, respectively.

A finite path in $\mathcal{A}$ is *successful* if its origin is in I and its end is in F. An infinite path p is *successful* if its origin is in I and if some state of F occurs infinitely often in p.

The set of finite (respectively infinite) words *recognized* by $\mathcal{A}$ is the set, denoted $L^+(\mathcal{A})$ (respectively $L^\omega(\mathcal{A})$), of the labels of all successful finite (respectively infinite) paths of $\mathcal{A}$. A set of finite (respectively infinite) words X is *recognizable* if there exists a finite Büchi automaton $\mathcal{A}$ such that $X = L^+(\mathcal{A})$ (respectively $X = L^\omega(\mathcal{A})$).

Example 4.2 Let $\mathcal{A}$ be the Büchi automaton obtained from example 4.1 by taking $I = \{1\}$ and $F = \{2\}$. Initial states are represented by an incoming arrow and final states by an arrow going out.

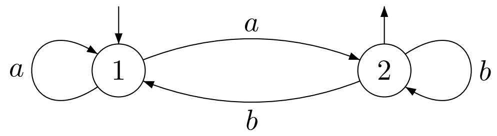

Figure 4.2: A Büchi automaton.

Then $L^+(\mathcal{A}) = a\{a,b\}^*$ is the set of all finite words whose first letter is an a, $L^\omega(\mathcal{A}) = a(a^*b)^\omega$ is the set of infinite words whose first letter is an a and containing an infinite number of b's.

Example 4.3 Let $\mathcal{A}$ be the Büchi automaton represented below. Then $L^\omega(\mathcal{A}) = A^*a^\omega$ is the set of all infinite words containing a finite number of b's.

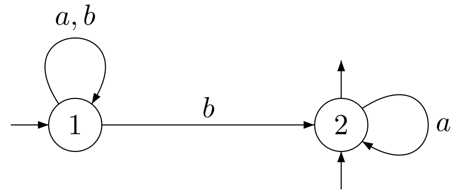

Figure 4.3: The automaton $\mathcal{A}$ recognizing A^*a^ω.

Let $\mathcal{A}'$ be the automaton obtained from $\mathcal{A}$ by taking only 1 as initial state. Then $L^\omega(\mathcal{A}') = A^*ba^\omega$.

Example 4.4 Let $\mathcal{A}$ be the Büchi automaton represented below

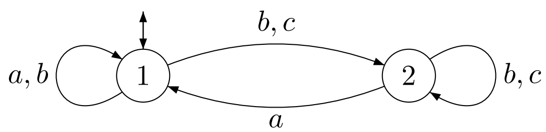

Figure 4.4:

Then $L^\omega(\mathcal{A}) = (a\{b,c\}^* \cup \{b\})^\omega$.

The relationship between rational and recognizable sets of finite words is given by the famous theorem of Kleene.

Theorem 4.1 *A subset of A^* is rational if and only if it is recognizable.*

The counterpart of Kleene's theorem for infinite words is due to Büchi [4].

Theorem 4.2 *A subset of $A^{\mathbb{N}}$ is ω-rational if and only if it is recognizable.*

Sketch of the proof.

(1) Let $\mathcal{A} = (Q, A, E, I, F)$ be a finite Büchi automaton. For $p, q \in Q$, denote by $L^+(E, p, q)$ the set of non empty finite words recognized by the automaton $(Q, A, E, \{p\}, \{q\})$. By Theorem 4.1, all these sets are rational. Furthermore, it is not difficult to see that

$$X = L^\omega(\mathcal{A}) = \bigcup_{i \in I} \bigcup_{f \in F} L^+(E, i, f)\bigl(L^+(E, f, f)\bigr)^\omega$$

It follows that $L^\omega(\mathcal{A})$ is ω-rational.

(2) It is easy to see that recognizable sets are closed under finite union. Indeed, the disjoint union of two automata $\mathcal{A}_1$ and $\mathcal{A}_2$ recognizes exactly the union of the sets recognized respectively by $\mathcal{A}_1$ and $\mathcal{A}_2$. In order to prove that every ω-rational set is recognizable, it remains to show that every set of the form XX'^ω, where X and X' are non-empty rational subsets of A^+, is recognizable. A small combinatorial argument shows that every rational set can be recognized by a *normal* automaton, that is, an automaton with exactly one initial state i, exactly one final state f and no transition starting from f or ending in i. Now, if $\mathcal{A}$ and $\mathcal{A}'$ are normal automata recognizing X and X', then the automaton obtained by identifying f, i' and f', as indicated in the figure below, recognizes XX'^ω:

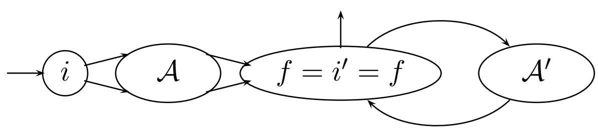

Figure 4.5: An automaton recognizing $X(X')^\omega$.

5 Deterministic Büchi automata

A Büchi automaton $\mathcal{A} = (Q, A, E, I, F)$ is said to be *deterministic* if I is a singleton and if E contains no pair of transitions of the form $(q, a, q_1), (q, a, q_2)$ with $q_1 \neq q_2$. In other words, given a state q and a letter a, there is at most one state q' such that $(q, a, q') \in E$. In particular, each word u is the label of at most one path starting from the initial state. A finite word u is accepted if the end of this unique path is a final state. And an infinite word is accepted if the unique path visits infinitely often a final state. It follows that an infinite word is accepted by $\mathcal{A}$ if and only if it has infinitely many prefixes accepted by $\mathcal{A}$. This leads to the following definition: for any subset L of A^*, put

$$\overrightarrow{L} = \{u \in A^{\mathbb{N}} \mid u \text{ has infinitely many prefixes in } L\}.$$

Then one can state

Proposition 5.1 *If $\mathcal{A}$ is a finite deterministic Büchi automaton, then* $L^\omega(\mathcal{A}) = \overrightarrow{L^+(\mathcal{A})}$.

Example 5.1 If $A = \{a, b\}$ and $L = a^*b$, then $\overrightarrow{L} = \emptyset$.

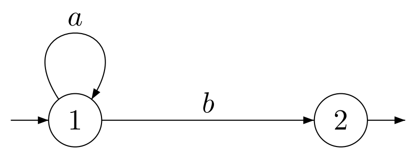

Figure 5.1: $L^+(\mathcal{A}) = a^*b$ and $L^\omega(\mathcal{A}) = \emptyset$.

Consider now the set L of all finite words of the form $a^n b$. Then $\overrightarrow{L} = (a^*b)^\omega$, the set of infinite words containing infinitely many b's.

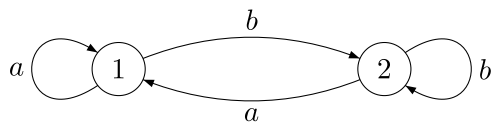

Figure 5.2: $L^+(\mathcal{A}) = (a^*b)^+ = \{a, b\}^*b$ and $L^\omega(\mathcal{A}) = (a^*b)^\omega$.

A set of infinite words is *deterministic* if it is recognized by a finite Büchi automaton. Landweber [14] has shown the existence of non deterministic recognizable sets.

Proposition 5.2 *The set $\{a, b\}^*a^\omega$ is not deterministic.*

Proof. Let $X = \{a, b\}^*a^\omega$ and suppose that $X = \overrightarrow{L}$ for some subset L of A^*. Then ba^ω has a prefix $u_1 = ba^{n_1}$ in L and similarly, $ba^{n_1}ba^\omega$ has a prefix $u_2 = ba^{n_1}ba^{n_2}$ in L, etc. so that the infinite word $u = ba^{n_1}ba^{n_2}ba^{n_3}\cdots$ has infinitely many prefixes in L. Thus $u \in \overrightarrow{L}$, a contradiction, since u has an infinite number of b's. $\square$

6 Muller automata

We have seen that non deterministic Büchi automata are not equivalent to deterministic ones. This lead Muller [22] to introduce a more satisfying definition.

A finite *Muller automaton* is a quintuple $\mathcal{A} = (Q, A, E, I, \mathcal{T})$ where

(1) (Q, A, E) is a finite automaton,

(2) I is a subset of Q, called the set of *initial* states.

(3) $\mathcal{T} = \{T_1, \ldots, T_n\}$ (the state table) is a set of subsets $T_1, \ldots, T_n$ of Q.

A path p is *successful* in $\mathcal{A}$ if $\text{Infinite}(p) \in \mathcal{T}$, where $\text{Infinite}(p)$ is the set of states that are visited infinitely often by p. The set $L^\omega(\mathcal{A})$ (also denoted $L^\omega(E, i, \mathcal{T})$) of infinite words accepted by $\mathcal{A}$ is the set of labels of all successful paths.

Example 6.1 Let $A = \{a, b\}$. The automaton below, with the state table $\mathcal{T} = \{\{2\}\}$, recognizes the set A^*b^ω of all infinite words containing a finite number of a's.

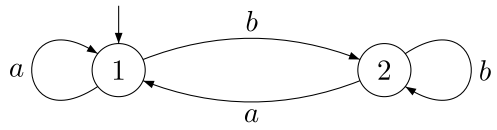

Figure 6.1: A Muller automaton.

The automaton $\mathcal{A}'$ obtained by taking as state table $\mathcal{T}' = \{\{1\}, \{2\}\}$, recognizes the set $A^*a^\omega \cup A^*b^\omega$.

A finite Muller automaton $\mathcal{A} = (Q, A, E, I, \mathcal{T})$ is *deterministic* if I is a singleton and if E contains no pair of transitions of the form $(q, a, q_1), (q, a, q_2)$ with $q_1 \neq q_2$. This time, deterministic automata are equivalent to non deterministic ones. The following result, mainly due to McNaughton [19], summarizes several equivalences.

Theorem 6.1 *Let X be a subset of $A^{\mathbb{N}}$. The following conditions are equivalent:*

(1) *X is ω-rational,*

(2) *X is recognized by a finite Muller automaton,*

(3) *X is recognized by a finite deterministic Muller automaton,*

(4) *X is recognized by a finite Büchi automaton,*

(5) *X is a boolean combination of sets recognized by a finite deterministic Büchi automaton.*

Sketch of the proof. (3) implies (2) is clear.

(2) implies (1). Let $\mathcal{A} = (Q, A, E, I, \mathcal{T})$ be a deterministic Muller automaton with $\mathcal{T} = \{T_1, \ldots, T_n\}$. Then, it is easy to see that

$$L^\omega(\mathcal{A}) = \bigcup_{1 \leqslant j \leqslant n} L^\omega(E, i, \{T_j\})$$

In other words, we may assume that the table $\mathcal{T}$ contains only one set of states $T = \{t_0, \ldots, t_{k-1}\}$. Let $X = L^+(E, i, t_0)$ and, for $0 \leqslant i \leqslant k-1$, let X_i be the set of labels of all finite paths from t_i to t_{i+1} visiting only states of $\mathcal{T}$ (where the indices i are calculated modulo k). Now a path p is successful if and only if it starts in i and $\text{Infinite}(p) = T$. In this case, p can be factorized as follows: a (finite) path from i to t_0, one from t_0 to t_1, one from t_1 to t_2, ..., one from t_{k-2} to t_{k-1}, one from t_{k-1} to t_0, one from t_0 to t_1, etc. Therefore

$$L^\omega(E, i, \{T\}) = X(X_0 X_1 \cdots X_{k-1})^\omega$$

and thus $L^\omega(E, i, \{T\})$ is ω-rational.

(1) implies (4) follows from Theorem 4.2.

(4) implies (3) is the more difficult part of the proof. We shall present in section 10 a semigroup theoretic proof of this result. The best known algorithm to convert a Büchi automaton into a deterministic Muller automaton is Safra's construction, which, given a n-state Büchi automaton, produces an equivalent deterministic Muller automaton with at most n^n states. The reader is referred to the original article of Safra [37] or to [30] for a presentation of this difficult algorithm.

Thus conditions (1) to (4) are equivalent. We conclude by showing that (2) implies (5) and (5) implies (1).

(2) implies (5). Let $\mathcal{A} = (Q, A, E, I, \mathcal{T})$ be a deterministic Muller automaton. As above, we may assume that $\mathcal{T} = \{T\}$. Then, for an infinite path p, the condition $\text{Infinite}(p) = T$ is equivalent to the conjunction of the two conditions

(1) p visits every $t \in T$ infinitely often,

(2) p visits every $t \notin T$ finitely often.

This can be converted into the following formula, which gives (5)

$$L^\omega(E, i, \{T\}) = \bigcap_{t \in T} L^\omega(E, i, t) \setminus \bigcap_{t \notin T} L^\omega(E, i, t)$$

(5) implies (1). By Theorem 4.2, every set recognized by a finite deterministic Büchi automaton is ω-rational. Furthermore, since (1) and (3) are equivalent, ω-rational sets are closed under boolean operations.

Corollary 6.2 *The recognizable sets of $A^{\mathbb{N}}$ are closed under finite boolean operations, morphisms and inverse morphisms between free semigroups.*

7 ω-semigroups

As it is shown elsewhere in this volume [33], one can use finite semigroups in place of automata to define the recognizable sets of finite words. It is possible to extend this idea to infinite words by replacing semigroups by ω-semigroups, which are, basically, algebras in which infinite products are defined. Although these algebras do not have a finitary signature, standard results on algebras still hold. In particular, A^∞ appears to be the free algebra on the set A and recognizable sets can be defined, as before, as the sets recognized by finite algebras. However, a problem arises since these finite algebras have an infinitary signature and thus are not finite objects. This problem can be solved by a Ramsey type argument showing that the structure of these finite algebras is totally determined by only three operations of finite signature. This defines a new type of algebras of finite signature, the Wilke algebras, that suffice to deal with infinite products. [1]

We now come to the precise definitions. An *ω-semigroup* is a two-sorted algebra $S = (S_f, S_\omega)$ equipped with the following operations:

(a) A binary operation defined on S_f and denoted multiplicatively,

(b) A mapping $S_f \times S_\omega \to S_\omega$, called *mixed product*, that associates to each pair $(s, t) \in S_f \times S_\omega$ an element of S_ω denoted st,

(c) A mapping $\pi : S_f^{\mathbb{N}} \to S_\omega$, called *infinite product*

These three operations satisfy the following properties :

(1) S_f, equipped with the binary operation, is a semigroup,

(2) for every $s, t \in S_f$ and for every $u \in S_\omega$, $s(tu) = (st)u$,

(3) for every increasing sequence $(k_n)_{n>0}$ and for all $(s_n)_{n\in\mathbb{N}} \in S_f^{\mathbb{N}}$,

$$\pi(s_0 s_1 \ \cdots \ s_{k_1-1}, s_{k_1} s_{k_1+1} \ \cdots \ s_{k_2-1}, \ \ldots) = \pi(s_0, s_1, s_2, \ldots)$$

(4) for every $s \in S_f$ and for every $(s_n)_{n\in\mathbb{N}} \in S_f^{\mathbb{N}}$

$$s\pi(s_0, s_1, s_2, \ldots) = \pi(s, s_0, s_1, s_2, \ldots)$$

Conditions (1) and (2) can be thought of as an extension of associativity. Conditions (3) and (4) show that one can replace the notation $\pi(s_0, s_1, s_2, \ldots)$ by the notation $s_0 s_1 s_2 \cdots$ without ambiguity . We shall use this simplified notation in the sequel. Intuitively, an ω-semigroup is a sort of semigroup in which infinite products are defined.

A ω-semigroup $S = (S_f, S_\omega)$ is *complete* if every element of S_ω can be written as an infinite product of elements of S_f. It is a harmless hypothesis to assume that all the ω-semigroups considered in this paper are complete.

Example 7.1 We denote by A^∞ the ω-semigroup $(A^+, A^{\mathbb{N}})$ equipped with the usual concatenation product.

Given two ω-semigroups $S = (S_f, S_\omega)$ and $T = (T_f, T_\omega)$, a *morphism of ω-semigroups* S is a pair $\varphi = (\varphi_f, \varphi_\omega)$ consisting of a semigroup morphism $\varphi_f : S_f \to T_f$ and of a mapping $\varphi_\omega : S_\omega \to T_\omega$ preserving the mixed product and the infinite product: for every sequence $(s_n)_{n\in\mathbb{N}}$ of elements of S_f,

$$\varphi_\omega(s_0 s_1 s_2 \cdots) = \varphi_f(s_0)\varphi_f(s_1)\varphi_f(s_2) \cdots$$

[1] Actually, the chronology is a little bit different. Ramsey type arguments have been used for a long time in semigroup theory [18, 40, 9], the Wilke algebras were introduced by Wilke in [45] under the name of *binoids* to clarify the approach of Arnold [1], Pécuchet [23] and Perrin [25] and the idea of using infinite products on semigroups came last [30].

and for every $s \in S_f$, $t \in S_\omega$,

$$\varphi_f(s)\varphi_\omega(t) = \varphi_\omega(st)$$

In the sequel, we shall often omit the subscripts, and use the simplified notation φ instead of φ_f and φ_ω.

Algebraic concepts like ω-subsemigroup, quotient and division are easily adapted to ω-semigroups. The semigroup A^+ is called the *free semigroup* on the set A because it satisfies the following property (which defines free objects in the general setting of category theory): every map from A into a semigroup S can be extended *in a unique way* into a semigroup morphism from A^+ into S. Similarly, it is not difficult to see that the free ω-semigroup on $(A, \emptyset)$ is the ω-semigroup A^∞.

A key result is that when S is finite, the infinite product is totally determined by the elements of the form $s^\omega = sss\cdots$, according to the following result of Wilke [46].

Theorem 7.1 *Let S_f be a finite semigroup and let S_ω be a finite set. Suppose that there exists a mixed product $S_f \times S_\omega \to S_\omega$ and a map from S_f into S_ω, denoted $s \to s^\omega$, satisfying, for every $s, t \in S_f$, the equations*

$$\begin{aligned} s(ts)^\omega &= (st)^\omega \\ (s^n)^\omega &= s^\omega \quad \textit{for every } n > 0 \end{aligned}$$

Then the pair $S = (S_f, S_\omega)$ can be equipped, in a unique way, with a structure of ω-semigroup such that for every $s \in S$, the product $sss\cdots$ is equal to s^ω.

This is a non trivial result, based on a consequence of Ramsey's theorem which is worth mentioning:

Theorem 7.2 *Let $\varphi : A^+ \to S$ be a morphism from A^+ into a finite semigroup S. For every infinite word $u \in A^{\mathbb{N}}$, there exist a pair (s, e) of elements of S such that $se = s$, $e^2 = e$, and a factorization $u = u_0u_1\cdots$ of u as a product of words of A^+ such that $\varphi(u_0) = s$ and $\varphi(u_n) = e$ for every $n > 0$.*

This motivates the following definition: a *linked pair* in a finite semigroup S is a pair $(s, e) \in S \times S$ such that e is idempotent and $se = s$. Two linked pairs (s, e) and (s', e') are *conjugate* (notation $(s, e) \sim (s', e')$) if there exist $x, y \in S^1$ such that $e = xy$, $e' = yx$ and $s' = sx$. Note that these conditions also imply $s = s'y$, since $s'y = sxy = se = s$.

A *Wilke algebra* is a two-sorted algebra (S_f, S_ω) equipped with three operations: an associative product on S_f, a mixed product $S_f \times S_\omega \to S_\omega$ and a map from S_f into S_ω, denoted $s \to s^\omega$, satisfying, for every $s, t \in S_f$, the equations

$$\begin{aligned} s(ts)^\omega &= (st)^\omega \\ (s^n)^\omega &= s^\omega \quad \text{for every } n > 0 \end{aligned}$$

A Wilke algebra is *complete* if every element of S_ω can be written as st^ω for some $s, t \in S_f$. Theorem 7.1 states that every finite ω-semigroup is equivalent with a finite Wilke algebra.

One can attach to every finite semigroup S a finite complete ω-semigroup $\bar{S} = (S, S_\omega)$, constructed as follows. Denote by π the exponent of S, that is, the smallest positive integer n such that s^n is idempotent for all $s \in S$. Let $[s, e]$ be the conjugacy class of a linked pair (s, e). One defines S_ω as the set of conjugacy classes of linked pairs of S. The pair $\bar{S}$ is equipped with a structure of Wilke algebra by setting, for all $[s, e] \in S_\omega$ and $t \in S$,

$$t[s, e] = [ts, e] \quad \text{and} \quad t^\omega = [t^\pi, t^\pi]$$

The definition is coherent, since if (s', e') and (s, e) are conjugate linked pairs, then (ts', e') and (ts, e) are conjugate linked pairs. One can prove that $\bar{S}$ is a complete Wilke algebra such that every semigroup morphism $\varphi : A^+ \to S$ can be extended into a morphism of ω-semigroups $\varphi : A^\infty \to \bar{S}$ defined by $\bar{\varphi}_f = \varphi$ and $\bar{\varphi}_\omega(u) = [s, e]$, where (s, e) is a linked pair associated with u.

The ω-semigroup $\bar{S}$ has the following universal property.

Proposition 7.3 *Let $T = (T_f, T_\omega)$ be a finite ω-semigroup and let S be a finite semigroup. For every semigroup morphism φ from S into T_f, there exists a unique morphism of ω-semigroups $\bar{\varphi} : \bar{S} \to T$ that extends φ. Furthermore, if $\varphi(S)$ generates T as an ω-semigroup, then $\bar{\varphi}$ is onto.*

In particular, every finite ω-semigroup is a quotient of an ω-semigroup of the form $\bar{S}$.

Corollary 7.4 *Every finite complete ω-semigroup $S = (S_f, S_\omega)$ is a quotient of $\bar{S}_f$.*

Example 7.2 Let $S = \{a, b\}$ be the finite semigroup given by the multiplication table $aa = a$, $ab = a$, $ba = b$ and $bb = b$. There are four linked pairs and two conjugacy classes: $(a, a) \sim (a, b)$ and $(b, a) \sim (b, b)$. Thus, $S_\omega = \{[a, a], [b, b]\}$ and the ω-power and the mixed product are given in the following tables

$$\begin{aligned} a^\omega &= [a, a] \qquad\qquad & b^\omega &= [b, b] \\ a[a, a] = a[b, b] &= [a, a] & b[a, a] = b[b, b] &= [b, b] \end{aligned}$$

Example 7.3 Let S be the five element Brandt semigroup: S is the semigroup with zero presented on the set $\{a, b\}$ by the relations $a^2 = 0$, $b^2 = 0$, $aba = a$ and $bab = b$. Thus $S = \{a, b, ab, ba, 0\}$ and the $\mathcal{D}$-class structure of S is represented in the figure below:

$$\begin{array}{|c|c|}\hline {}^{*}ab & a \\ \hline b & {}^{*}ba \\ \hline \end{array}$$

$$\begin{array}{|c|}\hline {}^{*}0 \\ \hline \end{array}$$

There are seven linked pairs and four conjugacy classes: $(0, ab) \sim (0, ba)$, $(a, ba) \sim (ab, ab)$, $(b, ab) \sim (ba, ba)$ and $(0, 0)$. Thus,

$$S_\omega = \{[ab, ab], [ba, ba], [0, ab], [0, 0]\}$$

and the ω-power and the mixed product are given in the following tables

$$(ab)^\omega = [ab, ab] \qquad (ba)^\omega = [ba, ba] \qquad a^\omega = b^\omega = [0, 0]$$

$$\begin{array}{rlll}
a[ab, ab] = [0, ab] & a[ba, ba] = [ab, ab] & a[0, ab] = [0, ab] & a[0, 0] = [0, 0] \\
b[ab, ab] = [ba, ba] & b[ba, ba] = [0, ab] & b[0, ab] = [0, ab] & b[0, 0] = [0, 0] \\
ab[ab, ab] = [ab, ab] & ab[ba, ba] = [0, ab] & ab[0, ab] = [0, ab] & ab[0, 0] = [0, 0] \\
ba[ab, ab] = [0, ab] & ba[ba, ba] = [ba, ba] & ba[0, ab] = [0, ab] & ba[0, 0] = [0, 0] \\
0[ab, ab] = [0, ab] & 0[ba, ba] = [0, ab] & 0[0, ab] = [0, ab] & 0[0, 0] = [0, 0]
\end{array}$$

A surjective morphism of ω-semigroups $\varphi : A^\infty \to S$ *recognizes a subset* X of $A^{\mathbb{N}}$ if there exists a subset P of S_ω such that $X = \varphi^{-1}(P)$. By extension, an ω-semigroup S recognizes X if there exists a surjective morphism of ω-semigroups $\varphi : A^\infty \to S$ that recognizes X. This definition can be extended to subsets of A^∞ as follows. A morphism of ω-semigroups $\varphi : A^\infty \to S$ *recognizes a subset* X of A^∞ if and only if $\varphi^{-1}\varphi(X) = X$, that is, if there exists a subset $P = (P_f, P_\omega)$ of (S_f, S_ω) such that $X \cap A^+ = \varphi_f^{-1}(P_f)$ and $X \cap A^\infty = \varphi_\omega^{-1}(P_\omega)$. A consequence of Corollary 7.4 is that every subset of A^∞ recognized by a finite ω-semigroup is recognized by an ω-semigroup of the form $\bar{S}$.

Proposition 7.5 *Let $S = (S_f, S_\omega)$ be a finite ω-semigroup recognizing a subset X of A^∞. Then $\bar{S}_f$ also recognizes X.*

Example 7.4 Let $S = (\{a, b\}, \{a^\omega, b^\omega\})$, where the finite product is given by the equalities $aa = a$, $ab = a$, $ba = b$, $bb = b$ and the infinite product is given by the rules $ax_1x_2\cdots = a^\omega$ and $bx_1x_2\cdots = b^\omega$. One recognizes here the ω-semigroup of Example 7.2. Let $\varphi : A^\infty \to S$ be the morphism of ω-semigroups defined by $\varphi(a) = a$ and $\varphi(b) = b$. Then $\varphi^{-1}(a^\omega) = aA^\omega$.

As for finite words, the following theorem holds.

Theorem 7.6 *A subset of $A^{\mathbb{N}}$ is recognizable if and only if it is recognized by a finite ω-semigroup.*

In other words, finite ω-semigroups can be converted into automata and vice versa. It is not difficult to see that a subset of $A^{\mathbb{N}}$ recognized by a finite ω-semigroup is ω-rational and thus, by Theorem 4.2, recognizable. Indeed, if $\varphi : A^\infty \to S$ recognizes X, then X is a finite union of sets of the form $\varphi^{-1}(s)\varphi^{-1}(e)^\omega$. The algorithm to pass from a finite ω-semigroup to a Büchi automaton presented in section 9 gives a direct proof of this result, without using Theorem 4.2. In the opposite direction, an algorithm to pass from Büchi automata to ω-semigroups is presented in section 8. Finally, the algorithm to pass from a finite ω-semigroup to a Muller automaton given in section 10 gives a proof of McNaughton's theorem.

8 From Büchi automata to ω-semigroups

We give in this section the construction to pass from a finite (Büchi) automaton to a finite ω-semigroup. This construction is much more involved than the corresponding construction for finite words.

Given a finite Büchi automaton $\mathcal{A} = (Q, A, E, I, F)$ recognizing a subset of X of $A^{\mathbb{N}}$, we would like to obtain a finite ω-semigroup recognizing X. Our construction makes use of the semiring $k = \{-\infty, 0, 1\}$ in which addition is the maximum for the ordering $-\infty < 0 < 1$ and multiplication, which extends the boolean addition, is given in the following table

	$-\infty$	0	1
$-\infty$	$-\infty$	$-\infty$	$-\infty$
0	$-\infty$	0	1
1	$-\infty$	1	1

Table 1: The multiplication table.

To each letter $a \in A$ is associated a matrix $\mu(a)$ with entries in k defined by

$$\mu(a)_{p,q} = \begin{cases} -\infty & \text{if } (p,a,q) \notin E \\ 0 & \text{if } (p,a,q) \in E \text{ and } p \notin F \text{ and } q \notin F \\ 1 & \text{if } (p,a,q) \in E \text{ and } (p \in F \text{ or } q \in F) \end{cases}$$

We have already used a similar technique to encode automata, but now we want to keep track of the visits of final states. We would like to extend μ into a morphism of ω-semigroups. It is easy to extend μ to a semigroup morphism from A^+ to the multiplicative semigroup of $Q \times Q$-matrices over

k. If u is a finite word, one gets

$$\mu(u)_{p,q} = \begin{cases} -\infty & \text{if there exists no path of label } u \text{ from } p \text{ to } q, \\ 1 & \text{if there exists a path from } p \text{ to } q \text{ with label } u \\ & \text{going through a final state,} \\ 0 & \text{if there exists a path from } p \text{ to } q \text{ with label } u \\ & \text{but no such path goes through a final state} \end{cases}$$

However, trouble arises when one tries to equip $k^{Q\times Q}$ with a structure of ω-semigroup. The solution consists in coding infinite paths not by square matrices, but by column matrices, in such a way that each coefficient $\mu(u)_p$ codes the existence of an infinite path of label u starting at p.

Let $S = (S_f, S_\omega)$ where $S_f = k^{Q\times Q}$ is the set of square matrices of size $\operatorname{Card} Q$ with entries in k and $S_\omega = k^Q$ is the set of column matrices with entries in $\{-\infty, 1\}$.

In order to define the operation ω on square matrices, we need the following definition. If s is a matrix of S_f, we call *infinite s-path starting at p* a sequence $p = p_0, p_1, \ldots$ of elements of Q such that, for $0 \leqslant i \leqslant n-1$, $s_{p_i,p_{i+1}} \neq -\infty$.

The s-path p is *successful* if $s_{p_i,p_{i+1}} = 1$ for an infinite number of coefficients. Then s^ω is the element of S_ω defined, for every $p \in Q$, by

$$s^\omega_p = \begin{cases} 1 & \text{if there exists a successful } s\text{-path of origin } p, \\ -\infty & \text{otherwise} \end{cases}$$

Note that the coefficients of this matrix can be effectively computed. Indeed, computing s^ω_p amounts to checking the existence of circuits containing a given edge in a finite graph. Then one can verify that S, equipped with these operations, is an ω-semigroup. The morphism μ can now be extended in a unique way as a morphism of ω-semigroups from A^∞ into S. Furthermore, we have the following result.

Proposition 8.1 *The morphism of ω-semigroups from A^∞ into S induced by μ recognizes the set $L^\omega(\mathcal{A})$.*

The ω-semigroup $\mu(A^\infty)$ is called the ω-semigroup *associated with* $\mathcal{A}$.

Example 8.1 Let $X = (a\{b,c\}^* \cup \{b\})^\omega$. This set is recognized by the Büchi automaton represented below:

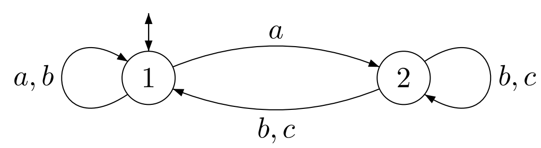

Figure 8.1:

The ω-semigroup associated with this automaton contains 9 elements

$$a = \begin{pmatrix} 1 & 1 \\ -\infty & -\infty \end{pmatrix} \qquad b = \begin{pmatrix} 1 & -\infty \\ 1 & 0 \end{pmatrix} \qquad c = \begin{pmatrix} -\infty & -\infty \\ 1 & 0 \end{pmatrix}$$

$$ba = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad ca = \begin{pmatrix} -\infty & -\infty \\ 1 & 1 \end{pmatrix} \qquad a^\omega = \begin{pmatrix} 1 \\ -\infty \end{pmatrix}$$

$$b^\omega = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad c^\omega = \begin{pmatrix} -\infty \\ -\infty \end{pmatrix} \qquad (ca)^\omega = \begin{pmatrix} -\infty \\ 1 \end{pmatrix}$$

It is defined by the following relations

$$\begin{array}{ccccc}
a^2 = a & ab = a & ac = a & b^2 = b & bc = c \\
cb = c & c^2 = c & (ba)^\omega = b^\omega & aa^\omega = a^\omega & ab^\omega = a^\omega \\
ac^\omega = c^\omega & a(ca)^\omega = a^\omega & ba^\omega = b^\omega & bb^\omega = b^\omega & bc^\omega = c^\omega \\
b(ca)^\omega = (ca)^\omega & ca^\omega = (ca)^\omega & cb^\omega = (ca)^\omega & cc^\omega = c^\omega & c(ca)^\omega = (ca)^\omega
\end{array}$$

9 From Wilke algebras to Büchi automata

Given a morphism of ω-semigroups recognizing a subset X of $A^{\mathbb{N}}$, we construct a Büchi automaton that recognizes X. By Proposition 7.5, we may suppose that the ω-semigroup is of the form $\bar{S} = (S, S_\omega)$. Thus, let $\varphi : A^\infty \to \bar{S}$ be a morphism of ω-semigroups recognizing X and let $P = \varphi(X)$. The construction of a Büchi automaton that recognizes X relies on the following result of Pécuchet [24] which gives, for each idempotent e of S, a representation of S by relations in S^1.

Lemma 9.1 *Let S be a finite semigroup and e an idempotent of S. The map $\varphi_e : S \to \mathbb{B}^{S^1 \times S^1}$ defined, for all $s \in S$ and for all $p, q \in S^1$ by*

$$\varphi_e(s)_{p,q} = \begin{cases} 1 & \textit{if } sq = p \textit{ or if } sq = pe \\ 0 & \textit{otherwise} \end{cases}$$

is a semigroup morphism.

Let us consider, for every idempotent e of S, the Büchi automaton

$$\begin{aligned} &\mathcal{A}_e = (S^1, A, E_e, I_e, \{1\}) \quad \text{with } I_e = \{s \in S \mid (s, e) \in P\} \text{ and} \\ &E_e = \{(p, a, q) \in S^1 \times A \times S^1 \mid \varphi(a)q = p \text{ or } \varphi(a)q = pe\} \end{aligned}$$

Then the Büchi automaton $\mathcal{A}$ that recognizes X is the disjoint union of the automata $\mathcal{A}_e$, where e ranges over the set of idempotents of S.

Example 9.1 Let $\bar{S} = (\{a, b\}, \{a^\omega, b^\omega\})$ be the ω-semigroup considered in Example 7.2 and let $\varphi : A^\infty \to \bar{S}$ be the morphism of ω-semigroups defined by $\varphi(a) = a$ and $\varphi(b) = b$. The automaton $\mathcal{A}$ is represented in the figure below.

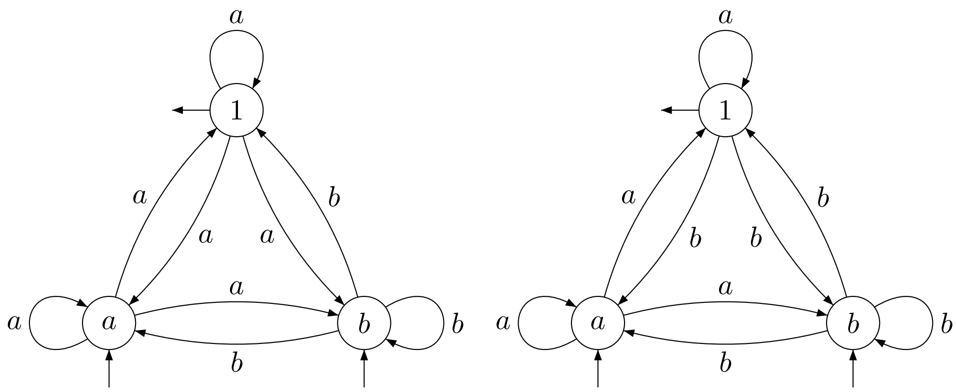

Figure 9.1: The automata $\mathcal{A}_a$ (on the left) and $\mathcal{A}_b$ (on the right).

10 From Wilke algebras to Muller automata

Given a semigroup morphism $\varphi : A^+ \to S$ recognizing a subset X of A^+, it is easy to build a deterministic automaton recognizing X. Denote by S^1 the monoid equal to S if S has an identity and to $S \cup \{1\}$ otherwise. Take the right representation of A^+ on S^1 defined by $s.a = s\varphi(a)$. This defines a deterministic automaton $\mathcal{A}_S = (S^1, A, E, \{1\}, P\}$, where $E = \{(s, a, s.a) \mid s \in S^1, a \in A\}$, which recognizes L. It is tempting to do the same construction to obtain a deterministic Muller automaton, given a finite ω-semigroup $\bar{S} = (S, S_\omega)$ recognizing a subset X of $A^{\mathbb{N}}$. B. Le Saec [15] has observed that such a construction is possible if the right stabilizers of S are bands. Given a finite semigroup S and an element s of S recall that the right stabilizer of s is the subsemigroup $\mathrm{Stab}(s)$ of S defined by $\mathrm{Stab}(s) = \{t \mid st = s\}$.

Theorem 10.1 *Let $\bar{S} = (S, S_\omega)$ be a finite ω-semigroup such that the right stabilizers of S are idempotent semigroups. Let $\varphi : A^\infty \to S$ be a morphism of ω-semigroups recognizing a subset X of $A^{\mathbb{N}}$. Then the automaton $\mathcal{A}_S = (S^1, A, E, \{1\}, \mathcal{T})$, where $E = \{(s, a, s.a) \mid s \in S^1, a \in A\}$ and $\mathcal{T} = \{\mathit{Infinite}(u) \mid u \in X\}$ is a Muller automaton recognizing X.*

Example 10.1 Let $\bar{S} = (\{a, b\}, \{a^\omega, b^\omega\})$ be the ω-semigroup considered in Examples 7.2 and 9.1 and let $\varphi : A^\infty \to \bar{S}$ be the morphism of ω-semigroups defined by $\varphi(a) = a$ and $\varphi(b) = b$. Let $P = \{a^\omega\}$, so that $\varphi^{-1}(a^\omega) = aA^\omega$. Then the transitions of the automaton $\mathcal{A}_S$ are represented below and its table is $\mathcal{T} = \{\{a\}\}$.

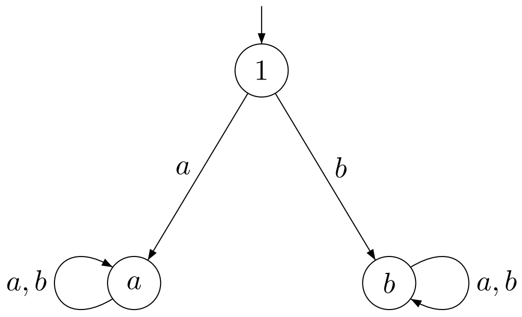

Figure 10.1: A deterministic Muller automaton recognizing aA^ω.

Unfortunately, the right stabilizer of an element of a finite semigroup is not always an idempotent semigroup. However, it is shown in [16, 17] that every finite semigroup is a quotient of a finite semigroup in which the right stabilizer of every element is an idempotent semigroup. In fact, a slightly stronger result holds.

Theorem 10.2 *Every finite semigroup is a quotient of a finite semigroup in which the right stabilizer of every element is a semigroup which satisfies the identities $x = x^2$ and $xy = xyx$.*

Therefore, one can associate with every morphism from A^∞ onto a finite ω-semigroup S a Muller automaton as follows: first compute a finite semigroup T whose right stabilizers are idempotent and commutative such that S_f is a quotient of T and then apply the result of Le Saec. This gives the proof of (4) implies (3) in Theorem 6.1. To sum up, we have obtained the following result.

Theorem 10.3 *Let X be a set of infinite words. The following conditions are equivalent:*

(1) *X is recognizable,*
(2) *X is recognized by a finite ω-semigroup,*
(3) *X is ω-rational.*

11 Syntactic ω-semigroup

Let $S = (S_f, S_\omega)$ be an ω-semigroup. A *congruence of ω-semigroup* on S is a pair $\sim\ = (\sim_f, \sim_\omega)$ where $\sim_f$ is a semigroup congruence on S_f and $\sim_\omega$ is an equivalence on S_ω such that

(1) for all $s, s' \in S_f$ and for all $t, t' \in S_\omega$, $s \sim_f s'$ and $t \sim_\omega t'$ imply $st \sim_\omega s't'$
(2) for all infinite sequences $s_0, s_1, s_2, \ldots$ and $s'_0, s'_1, s'_2, \ldots$ of elements of S_f such that $s_i \sim_f s'_i$ for all i, $s_0 s_1 \cdots \sim_\omega s'_0 s'_1 \cdots$

If S is a finite ω-semigroup, then one may replace (2) by the weaker condition (2′)

(2′) for all $s, s' \in S_f$, $s \sim_f s'$ implies $s_\omega \sim_\omega s'_\omega$.

In the sequel, we shall often omit the subscripts, and use the simplified notation $\sim$ instead of $\sim_f$ and $\sim_\omega$.

Let X be a subset of A^∞. A congruence of ω-semigroup $\sim$ on A^∞ recognizes X if the morphism of ω-semigroup $\varphi : A^\infty \to A^\infty/{\sim}$ recognizes X. Contrary to the case of finite words, the lower bound of the congruences that recognize X does not always recognize X. If this lower bound still recognizes X, it is called the *syntactic congruence* of X and is denoted $\sim_X$.

Arnold has shown that the syntactic congruence always exists if X is recognized by a finite ω-semigroup [1]. The reason is that a finite ω-semigroup can be considered as a finite Wilke algebra and that syntactic congruences can be defined on Wilke algebras as on any algebra of finite signature. More precisely, a *congruence* on a Wilke algebra S is a pair $\sim\ = (\sim_f, \sim_\omega)$ where $\sim_f$ is a semigroup congruence on S_f and $\sim_\omega$ is an equivalence relation on S_ω that is compatible with the ω-power and the mixed product. In practice, it is convenient to omit the subscripts f and ω and write $\sim$ for both $\sim_f$ and $\sim_\omega$.

A congruence $\sim$ on S *saturates* a subset $P = (P_f, P_\omega)$ of S if for all $x, y \in S_f$ (resp. S_ω), $x \sim y$ and $x \in P_f$ (resp. P_ω) implies $y \in P_f$ (resp. P_ω). Now the lower bound of the congruences that saturate P also saturates P and is called the *syntactic congruence* of P. It is the congruence $\sim_P$ defined on S_f by $u \sim_P v$ if and only if, for every $x, y \in S_f^1$ and for every $z \in S_f$,

$$\begin{aligned} xuyz^\omega \in P_\omega &\iff xvyz^\omega \in P_\omega \\ x(uy)^\omega \in P_\omega &\iff x(vy)^\omega \in P_\omega \end{aligned}$$

and on S_ω by $u \sim_P v$ if and only if, for every $x \in S_f^1$,

$$xu \in P_\omega \iff xv \in P_\omega \tag{11.1}$$

The *syntactic Wilke algebra* of P is the quotient of S by the syntactic congruence of P.

Now, to compute the syntactic ω-semigroup of a set X recognized by a finite Büchi automaton $\mathcal{A}$, one first computes the finite ω-semigroup S associated with $\mathcal{A}$ and the image P of X in S and then one computes the syntactic Wilke algebra of P in S. This provides an algorithm to compute the syntactic ω-semigroup of a recognizable set.

The syntactic congruence $\sim_X$ of a recognizable set X of A^∞ can also be defined directly as follows. On A^+, $u \sim_X v$ if and only if, for every $x, y \in A^*$ and for every $z \in A^+$,

$$\begin{aligned} xuyz^\omega \in X &\iff xvyz^\omega \in X \\ x(uy)^\omega \in X &\iff x(vy)^\omega \in X \end{aligned} \tag{11.2}$$

and on $A^{\mathbb{N}}$, $u \sim_X v$ if and only if, for every $x \in A^*$,

$$xu \in X \Longleftrightarrow xv \in X \tag{11.3}$$

The *syntactic ω-semigroup* of X is the quotient of A^∞ under the congruence of ω-semigroup $\sim_X$. This is also the minimal (with respect to the quotient ordering) complete finite ω-semigroup recognizing X.

Example 11.1 We come back to the example 8.1. The set $X = (a\{b,c\}^* \cup \{b\})^\omega$ is recognized by the ω-semigroup $S = \{a, b, c, ba, ca, a^\omega, b^\omega, (ca)^\omega, 0\}$, defined by the following relations:

$$\begin{array}{ccccc}
a^2 = a & ab = a & ac = a & b^2 = b & bc = c \\
cb = c & c^2 = c & c^\omega = 0 & (ba)^\omega = b^\omega & aa^\omega = a^\omega \\
ab^\omega = a^\omega & a(ca)^\omega = a^\omega & ba^\omega = b^\omega & bb^\omega = b^\omega & b(ca)^\omega = (ca)^\omega \\
ca^\omega = (ca)^\omega & cb^\omega = (ca)^\omega & c(ca)^\omega = (ca)^\omega & &
\end{array}$$

Since $P = \{a^\omega, b^\omega\}$, the congruence $\sim_P$ is defined by

$$a \sim_P ba \qquad \text{et} \qquad a^\omega \sim_P b^\omega$$

Therefore, the syntactic ω-semigroup is $S(X) = \{a, b, c, ca, a^\omega, (ca)^\omega, 0\}$, with the following relations

$$\begin{array}{ccccc}
a^2 = a & ab = a & ac = a & ba = a & b^2 = b \\
bc = c & cb = c & c^2 = c & b^\omega = a^\omega & c^\omega = 0 \\
aa^\omega = a^\omega & a(ca)^\omega = a^\omega & ba^\omega = a^\omega & b(ca)^\omega = (ca)^\omega & ca^\omega = (ca)^\omega \\
c(ca)^\omega = (ca)^\omega & & & &
\end{array}$$

12 Varieties

So far, we have extended to infinite words the notions of finite automata, of a semigroup recognizing a set of words and finally of syntactic semigroup. The next natural step in this process is to extend the variety theorem. Pécuchet [24, 23] made a first attempt in this direction, but the result of Wilke [45] is more satisfactory.

A *variety of finite Wilke algebras* is a class of finite Wilke algebras closed under the taking of finite direct products, quotients and complete subalgebras.

Let $X \subset A^\infty$ and let $u \in A^+$. We set

$$\begin{aligned}
u^{-1}X &= \{v \in A^\infty \mid uv \in X\} \\
Xu^{-\omega} &= \{v \in A^+ \mid (vu)^\omega \in X\} \\
Xu^{-1} &= \{v \in A^+ \mid vu \in X\}
\end{aligned}$$

An ∞-*variety* $\mathcal{V}$ associates with every alphabet A a set $\mathcal{V}(A^\infty)$ of recognizable sets of A^∞ such that

(1) for every alphabet A, $\mathcal{V}(A^\infty)$ contains $\emptyset$, A^+, $A^{\mathbb{N}}$ and A^∞ and is closed under finite boolean operations (finite union and complement),

(2) for every semigroup morphism $\varphi : A^+ \to B^+$, $X \in \mathcal{V}(B^\infty)$ implies $\varphi^{-1}(X) \in \mathcal{V}(A^\infty)$,

(3) If $X \in \mathcal{V}(A^\infty)$ and $u \in A^*$, then $u^{-1}X \in \mathcal{V}(A^\infty)$, $Xu^{-1} \in \mathcal{V}(A^\infty)$ and $Xu^{-\omega} \in \mathcal{V}(A^\infty)$.

It is important to notice that one works with A^∞ (and not with $A^{\mathbb{N}}$) in this definition. In other words, the elements of a ∞-variety are sets of finite or infinite words.

If $\mathbf{V}$ is a variety of finite Wilke algebras, we denote by $\mathcal{V}(A^\infty)$ the set of recognizable sets of A^∞ whose syntactic ω-semigroup belongs to $\mathbf{V}$. This is also the set of sets of A^∞ recognized by an ω-semigroup of $\mathbf{V}$.

An ∞-*class* of recognizable sets is a correspondence which associates with every finite alphabet A, a set $\mathcal{C}(A^\infty)$ of recognizable sets of A^∞. In particular, the correspondence $\mathbf{V} \to \mathcal{V}$ associates with every variety of finite Wilke algebras a ∞-class of recognizable sets.

The theorem of Wilke, which extends the corresponding theorem of Eilenberg, can now be stated as follows.

Theorem 12.1 *The correspondence $\mathbf{V} \to \mathcal{V}$ defines a bijection between the varieties of finite Wilke algebras and the ∞-varieties.*

We conclude by giving four examples of correspondence between varietiesof finite Wilke algebras and ∞-varieties. The first and the second one extend to infinite words well known results on finite words (the characterization of star-free and locally testable languages), but the last two examples are less familiar since they concern two topological classes.

We refer to our article [33] in this volume for the definitions of the star-free and locally testable languages. The set of *star-free subsets of* A^∞ is the smallest set $\mathcal{S}$ of subsets of A^∞ containing the star-free languages of A^+ which is closed under finite boolean operations and such that if X is a star-free subset of A^+ and $Y \in \mathcal{S}$, then $XY \in \mathcal{S}$. Equivalent characterizations of the star-free subsets of $A^{\mathbb{N}}$ were proposed by Thomas [41, 42].

Theorem 12.2 *Let X be a subset of $A^{\mathbb{N}}$. The following conditions are equivalent:*

(1) *X is star-free,*

(2) *X is a finite union of sets of the form UV^ω where U and V are star-free subsets of A^+ and $V^2 \subset V$,*

(3) *X is a boolean combination of sets of the form $\overrightarrow{L}$ where L is a star-free language.*

The syntactic characterization is given in [26]. A finite ω-semigroup (S_f, S_ω) is aperiodic if the semigroup S_f is aperiodic.

Theorem 12.3 *A recognizable set of A^∞ is star-free if and only if its syntactic ω-semigroup is aperiodic.*

It follows in particular that the star-free sets form a ∞-variety. Similar results hold for the locally testable sets. A subset of A^∞ is *locally testable* if and only if it is a finite boolean combination of sets of the form uA^*, A^*u, A^*uA^*, uA^ω, A^*uA^ω and $A^*(uA^*)^\omega$, for all $u \in A^+$. The following results are due to Pécuchet [24, 23]

Theorem 12.4 *Let X be a subset of $A^\mathbb{N}$. The following conditions are equivalent:*

(1) *X is locally testable,*

(2) *X is a finite union of sets of the form UV^ω where U and V are locally testable subsets of A^+ and $V^2 \subset V$,*

(3) *X is a boolean combination of sets of the form $\overrightarrow{L}$ where L is a locally testable language.*

A finite ω-semigroup (S_f, S_ω) is locally idempotent and commutative if the semigroup S_f is locally idempotent and commutative.

Theorem 12.5 *A recognizable set of A^∞ is locally testable if and only if its syntactic ω-semigroup is locally idempotent and commutative.*

One can define a topology on $A^\mathbb{N}$ by considering A as a discrete space and by taking the product topology. The open sets of $A^\mathbb{N}$ are the sets of the form $XA^\mathbb{N}$ for some $X \subset A^+$. A set is closed if its complement is open. The sets of Δ_1 are the clopen sets, sets which are at the same time closed and open. The sets of Δ_2 are at the same time countable unions of closed sets and countable intersection of open sets. One can show that the recognizable sets of Δ_2 are the sets which are accepted by a deterministic Büchi automaton as well as their complement. The characterization of these two classes in terms of ω-semigroups is due to Wilke [46]. In these statements, x^π denotes the unique idempotent of the semigroup generated by x.

Theorem 12.6 *The class Δ_1 form a ∞-variety. The corresponding variety of finite Wilke algebras is the class of algebras defined by the equation $x^\pi y z^\omega = x^\pi y' z'^\omega$.*

Theorem 12.7 *The class Δ_2 form a ∞-variety. The corresponding variety of finite Wilke algebras is the class of algebras defined by the equation $(x^\pi y^\pi)^\pi x^\omega = (x^\pi y^\pi)^\pi y^\omega$.*

13 Conclusion

Most known results on automata on finite words have found their counterpart on infinite words. However, the complexity of the solutions is apparently increased by an order of magnitude. This is true for instance for the equivalence between deterministic and non deterministic automata and for the variety theorem. Furthermore, several questions which are solved for finite words are still open for infinite words, but one can be reasonably optimistic about them.

Acknowledgements

The authors would like to thank Olivier Carton and Pascal Weil for their comments on a preliminary version of this article.

References

[1] A. Arnold, A syntactic congruence for rational ω-languages, *Theoret. Comput. Sci.* **39**,2-3 (1985), 333–335.

[2] D. Beauquier and J.-E. Pin, Factors of words, in *Automata, languages and programming (Stresa, 1989)*, pp. 63–79, Springer, Berlin, 1989.

[3] D. Beauquier and J.-E. Pin, Languages and scanners, *Theoret. Comput. Sci.* **84** (1991), 3–21.

[4] J. R. Büchi, Weak second-order arithmetic and finite automata, *Z. Math. Logik und Grundl. Math.* **6** (1960), 66–92.

[5] J. R. Büchi, On a decision method in restricted second order arithmetic, in *Logic, Methodology and Philosophy of Science (Proc. 1960 Internat. Congr .)*, pp. 1–11, Stanford Univ. Press, Stanford, Calif., 1962.

[6] J. R. Büchi and L. H. Landweber, Definability in the monadic second-order theory of successor, *J. Symbolic Logic* **34** (1969), 166–170.

[7] S. Eilenberg, *Automata, languages, and machines. Vol. A*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58.

[8] S. Eilenberg, *Automata, languages, and machines. Vol. B*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. With two chapters ("Depth decomposition theorem" and "Complexity of

semigroups and morphisms") by Bret Tilson, Pure and Applied Mathematics, Vol. 59.

[9] J. Justin and G. Pirillo, On a natural extension of Jacob's ranks, *J. Combin. Theory Ser. A* **43**,2 (1986), 205–218.

[10] S. C. Kleene, Representation of events in nerve nets and finite automata, in *Automata studies*, pp. 3–41, Princeton University Press, Princeton, N. J., 1956. Annals of mathematics studies, no. 34.

[11] R. E. Ladner, Application of model theoretic games to discrete linear orders and finite automata, *Information and Control* **33**,4 (1977), 281–303.

[12] G. Lallement, *Semigroups and combinatorial applications*, John Wiley & Sons, New York-Chichester-Brisbane, 1979. Pure and Applied Mathematics, A Wiley-Interscience Publication.

[13] L. H. Landweber, Finite state games - A solvability algorithm for restricted second-order arithmetic, *Notices Amer. Math. Soc.* **14** (1967), 129–130.

[14] L. H. Landweber, Decision problems for ω-automata, *Math. Systems Theory* **3** (1969), 376–384.

[15] B. Le Saec, Saturating right congruences, *RAIRO Inform. Théor. Appl.* **24**,6 (1990), 545–559.

[16] B. Le Saëc, J.-E. Pin and P. Weil, A purely algebraic proof of McNaughton's theorem on infinite words, in *Foundations of software technology and theoretical computer science (New Delhi, 1991)*, pp. 141–151, Springer, Berlin, 1991.

[17] B. Le Saëc, J.-E. Pin and P. Weil, Semigroups with idempotent stabilizers and applications to automata theory, *Internat. J. Algebra Comput.* **1**,3 (1991), 291–314.

[18] M. Lothaire, *Combinatorics on words*, Cambridge University Press, Cambridge, 1997. With a foreword by Roger Lyndon and a preface by Dominique Perrin, Corrected reprint of the 1983 original, with a new preface by Perrin.

[19] R. McNaughton, Testing and generating infinite sequences by a finite automaton, *Information and Control* **9** (1966), 521–530.

[20] R. McNaughton, Algebraic decision procedures for local testability, *Math. Systems Theory* **8**,1 (1974), 60–76.

[21] R. McNaughton and S. Papert, *Counter-free automata*, The M.I.T. Press, Cambridge, Mass.-London, 1971. With an appendix by William Henneman, M.I.T. Research Monograph, No. 65.

[22] D. E. Muller, Infinite sequences and finite machines, in *Switching Theory and Logical Design, Proc. Fourth Annual Symp. IEEE*, pp. 3–16, IEEE, 1963.

[23] J.-P. Pécuchet, Étude syntaxique des parties reconnaissables de mots infinis, in *Automata, languages and programming (Rennes, 1986)*, pp. 294–303, Springer, Berlin, 1986.

[24] J.-P. Pécuchet, Variétés de semigroupes et mots infinis, in *STACS 86*, B. Monien and G. Vidal-Naquet (ed.), pp. 180–191, *Lecture Notes in Comput. Sci.* vol. 210, Springer Verlag, Berlin, Heidelberg, New York, 1986.

[25] D. Perrin, Variétés de semigroupes et mots infinis, *C. R. Acad. Sci. Paris Sér. I Math.* **295**,10 (1982), 595–598.

[26] D. Perrin, Recent results on automata and infinite words, in *Mathematical foundations of computer science, 1984 (Prague, 1984)*, pp. 134–148, Springer, Berlin, 1984.

[27] D. Perrin, An introduction to finite automata on infinite words, in *Automata on infinite words (Le Mont-Dore, 1984)*, pp. 2–17, Springer, Berlin, 1985.

[28] D. Perrin, Finite automata, in *Handbook of theoretical computer science, Vol. B*, pp. 1–57, Elsevier, Amsterdam, 1990.

[29] D. Perrin and J.-E. Pin, First-order logic and star-free sets, *J. Comput. System Sci.* **32**,3 (1986), 393–406.

[30] D. Perrin and J.-E. Pin, Mots infinis, Rap. Tech., LITP, 1993. to appear (LITP report 93–40).

[31] D. Perrin and J.-E. Pin, Semigroups and automata on infinite words, in *Semigroups, formal languages and groups (York, 1993)*, pp. 49–72, Kluwer Acad. Publ., Dordrecht, 1995.

[32] D. Perrin and P. E. Schupp, Automata on the integers, recurrence, distinguishability and the equivalence and decidability of monadic theories, in *Proc. 1st IEEE Symp. on Logic in Computer Science*, pp. 301–304, IEEE, 1986.

[33] J.-E. Pin, Finite semigroups and recognizable languages: an introduction. this volume.

[34] J.-E. Pin, Logic, semigroups and automata on words. to appear.

[35] J.-E. Pin, *Varieties of formal languages*, Plenum Publishing Corp., New York, 1986. With a preface by M.-P. Schützenberger, Translated from the French by A. Howie.

[36] F. D. Ramsey, On a problem of formal logic, *Proc. of the London Math. Soc.* **30** (1929), 338–384.

[37] S. Safra, On the complexity of the ω-automata, in *Proc. 29th Ann. IEEE Symp. on Foundations of Computer Science*, pp. 319–327, IEEE, 1988.

[38] M.-P. Schützenberger, On finite monoids having only trivial subgroups, *Information and Control* **8** (1965), 190–194.

[39] I. Simon, Piecewise testable events, in *Proc. 2nd GI Conf.*, H. Brackage (ed.), pp. 214–222, *Lecture Notes in Comp. Sci.* vol. 33, Springer Verlag, Berlin, Heidelberg, New York, 1975.

[40] I. Simon, Word Ramsey theorems, in *Graph theory and combinatorics (Cambridge, 1983)*, pp. 283–291, Academic Press, London, 1984.

[41] W. Thomas, Star-free regular sets of ω-sequences, *Inform. and Control* **42**,2 (1979), 148–156.

[42] W. Thomas, A combinatorial approach to the theory of ω-automata, *Inform. and Control* **48**,3 (1981), 261–283.

[43] W. Thomas, Classifying regular events in symbolic logic, *J. Comput. System Sci.* **25**,3 (1982), 360–376.

[44] W. Thomas, Automata on infinite objects, in *Handbook of Theoretical Computer Science*, J. van Leeuwen (ed.), vol. vol. B, Formal models and semantics, pp. 135–191, Elsevier, 1990.

[45] T. Wilke, An Eilenberg theorem for ∞-languages, in *Automata, Languages and Programming*, pp. 588–599, *Lecture Notes in Computer Sci.* vol. 510, Springer Verlag, Berlin, Heidelberg, New York, 1991.

[46] T. Wilke, An algebraic theory for regular languages of finite and infinite words, *Int. J. Alg. Comput.* **3** (1993), 447–489.

[47] T. Wilke, Locally threshold testable languages of infinite words, in *STACS 93*, P. Enjalbert, A. Finkel and K. Wagner (ed.), pp. 607–616, *Lecture Notes in Comp. Sci.* vol. 665, Springer Verlag, Berlin, Heidelberg, New York, 1993.

[48] J. Wojciechowski, Classes of transfinite sequences accepted by finite automata, *Fundamenta Informaticae* **7** (1984), 191–223.

[49] J. Wojciechowski, Finite automata on transfinite sequences and regular expressions, *Fundamenta Informaticae* **8** (1985), 379–396.

[50] J. Wojciechowski, The ordinals less than ω^ω are definable by finite automata, in *Algebra, combinatorics and logic in computer science, Vol. I, II (Győr, 1983)*, pp. 871–887, North-Holland, Amsterdam, 1986.