

# Complementation of Rational sets on Scattered Linear Orderings of Finite rank

Olivier Carton<sup>1</sup> and Chloé Rispal<sup>2</sup>

<sup>1</sup> LIAFA, Université Paris 7, 2, place Jussieu, F-75251 Paris cedex 05, France,  
Olivier.Carton@liafa.jussieu.fr,

<sup>2</sup> IGM, Université de Marne-la-Vallée, 5 boulevard Descartes, F-77454  
Marne-la-Vallée Cedex 2, France,  
chloe.rispal@univ-mlv.fr

**Abstract.** In a preceding paper (Bruyère and Carton, automata on linear orderings, MFCS'01), automata have been introduced for words indexed by linear orderings. These automata are a generalization of automata for finite, infinite, bi-infinite and even transfinite words studied by Büchi. Kleene's theorem has been generalized to these words. We show that deterministic automata do not have the same expressive power. Despite this negative result, we prove that rational sets of words of finite ranks are closed under complementation.

## 1 Introduction

Automata were first introduced by Kleene who showed that they have the same expressive power as rational expressions [13]. Since then, many extensions of this deep result have been proved. Different kinds of structures have been considered like infinite words [7, 14], bi-infinite words [11, 15] and transfinite words [9, 10, 22], finite and infinite trees [18], finite and infinite traces, pictures, *etc.*

In [2, 3], have been introduced automata that accept linearly-ordered structures. These automata are a simple and natural generalization of usual automata with additional limit transitions of the form  $P \rightarrow q$  and  $q \rightarrow P$  where  $P$  is subset of states. They allow to treat in the same framework finite, infinite words, bi-infinite words and transfinite words. These automata were proved to be equivalent to some rational expressions when the orderings are restricted to scattered orderings. Recall that scattered orderings are those orderings which do not contain a dense sub-ordering like  $\mathbb{Q}$ . They include the ordinals and their mirrors.

One main property of rational sets is the closure under complementation. It means that for any automaton  $\mathcal{A}$ , there is another automaton  $\mathcal{B}$  accepting exactly the structures that are not accepted by  $\mathcal{A}$ . This property holds for almost all structures: finite and infinite words, finite and infinite trees and even for transfinite words on ordinals.

This property is important both from the practical and the theoretical point of view. It means that the class of rational sets forms an effective boolean algebra. It is used whenever some logic is translated into automata. For instance, in both

proofs of the decidability of the monadic second-order theory of the integers by Büchi [8] and the decidability of the monadic second-order theory of the infinite binary tree by Rabin [18], the closure under complementation of automata is the key property. It is well known that automata have the same expressive power as the monadic second order theory on many structures like finite, infinite and transfinite words and trees. A nice result would be to extend this equivalence to linear orderings. Proving the closure under complementation is one step towards this result.

In [3], the closure under complementation was left open. In this paper, we address this problem and we solve it for a subclass of scattered linear orderings. Namely, we prove that rational sets of words on scattered orderings of finite ranks are closed under complementation. Recall that Hausdorff's result [12] states that scattered orderings can be obtained from the finite orderings by repetitive applications of  $\omega$ -sums and  $-\omega$ -sums (see Theorem 1). The rank of a scattered linear ordering is the number of nested  $\omega$ -sums and  $-\omega$ -sums needed to obtain it. The ranks of all countable scattered linear orderings range over all countable ordinals. It can be seen as a measure of its complexity. For instance,  $\omega$  and  $\zeta$  are scattered orderings of rank 1. Our result generalizes both the complementation of infinite and bi-infinite words. The class of scattered orderings of finite rank includes ordinals smaller than  $\omega^\omega$ . Therefore, our result holds for sets of transfinite words studied by Choueka [10].

The classical method to get an automaton for the complement of a set of finite words accepted by an automaton  $\mathcal{A}$  is through determinization [1]. Another method uses algebraic objects like semigroups [17]. The determinization method can still be used for infinite words but it becomes more involved [21, 4]. This method has been pushed further by Büchi for countable transfinite words but it is then very complex [9]. The algebraic method can also be extended to ordinals [5, 6]. In our case, this method can not be applied since automata can not be made deterministic. In this paper, we give an example of a rational set of words that cannot be accepted by a deterministic automaton. Therefore, we use another method which was introduced by Büchi for infinite words. It is based on an equivalence relation on words whose classes are shown to be rational.

The paper is organized as follows. In Section 2, we introduce words indexed by linear orderings and recall the Hausdorff characterization of countable scattered linear orderings. Then rational sets of words are defined from rational operators and automata in section 3. We finally prove in section 4 that rational sets of words indexed by countable scattered linear orderings of finite ranks are closed under complementation.

## 2 Words on linear orderings

In this section, we recall some definitions and operations on linear orderings but we refer the reader to [20] for a complete introduction to linear orderings. We give the Hausdorff's characterization of countable scattered linear orderings and introduce words indexed by linear orderings.

Let  $J$  be a set equipped with an order  $<$ . The ordering  $J$  is *linear* if for any  $j$  and  $k$  in  $J$ , either  $j < k$  or  $k < j$ . A linear ordering  $J$  is *dense* if for any  $j$  and  $k$  in  $J$  such that  $j < k$ , there exists an element  $i$  of  $J$  such that  $j < i < k$ . It is *scattered* if it contains no dense subordering. The ordering  $\omega$  of natural integers and the ordering  $\zeta$  of relative integers are scattered. More generally, ordinals are scattered orderings.

Let  $A$  be a finite alphabet. A *word*  $x = (a_j)_{j \in J}$  indexed by a linear ordering  $J$  is a function from  $J$  to  $A$ .  $J$  is called the *length* of  $x$ . For instance  $\omega$  is the length of right-infinite words  $a_0 a_1 \dots$  and  $\zeta$  is the length of bi-infinite words  $\dots a_{-1} a_0 a_1 \dots$ .

In order to define the rank of scattered linear orderings, we recall operators.

## 2.1 Operations on linear orderings

For any linear ordering  $J$ , we denote by  $-J$  the backward linear ordering that is the set  $J$  equipped with the reverse ordering. For instance,  $-\omega$  is the linear ordering of negative integers.

The sum  $J + K$  of two linear orderings is the set  $J \cup K$  equipped with the ordering  $<$  extending the orderings of  $J$  and  $K$  by setting  $j < k$  for any  $j \in J$  and  $k \in K$ . For instance,  $\zeta = -\omega + \omega$ . Formally, the *sum*  $\sum_{j \in J} K_j$  is the set of all pairs  $(k, j)$  such that  $k \in K_j$  equipped with the ordering defined by  $(k_1, j_1) < (k_2, j_2)$  if and only if  $j_1 < j_2$  or  $(j_1 = j_2$  and  $k_1 < k_2$  in  $K_{j_1}$ ).

The sum of linear orderings helps to define the lengths of the products of words. Let  $J$  be a linear ordering and let  $(x_j)_{j \in J}$  be words of respective length  $K_j$  for any  $j \in J$ . The word  $x = \prod_{j \in J} x_j$  obtained by concatenation of the words  $x_j$  with respect to the ordering on  $J$  is of length  $L = \sum_{j \in J} K_j$ . We call *J-product* a product indexed by the ordering  $J$ . For instance, the  $\omega$ -product of the word  $a^\omega$  is the word  $(a^\omega)^\omega$  of length  $\sum_\omega \omega$ . The sequence  $(x_j)_{j \in J}$  of words is a *J-factorization* of the word  $x = \prod_{j \in J} x_j$ .

## 2.2 Construction of countable scattered linear orderings

Countable scattered linear orderings are defined through a forbidden pattern, namely that they do not contain a dense subordering. Hausdorff's theorem states that they can be constructed from finite orderings.

We denote by  $\mathcal{N}$  the subclass of finite linear orderings,  $\mathcal{O}$  the class of countable ordinals and  $\mathcal{S}$  the class of countable scattered linear orderings.

**Theorem 1.** [12] *A countable linear ordering  $J$  is scattered if and only if  $J$  belongs to  $\bigcup_{\alpha \in \mathcal{O}} V_\alpha$  where the classes  $V_\alpha$  are inductively defined by:*

1.  $V_0 = \{0, 1\}$

$$2. V_\alpha = \left\{ \sum_{j \in J} K_j \mid J \in \mathcal{N} \cup \{\omega, -\omega, \zeta\} \text{ and } K_j \in \bigcup_{\beta < \alpha} V_\beta \right\}.$$

where  $\mathbf{0}$  and  $\mathbf{1}$  are respectively the orderings of zero and one element.

Intuitively, the rank of a linear ordering is the maximum number of nested  $\omega$  and  $-\omega$ . It is linked to its Hausdorff's class. For instance the orderings  $\omega$  of rank 1 and  $\omega^2$  of rank 2 belong respectively to  $V_1$  and  $V_2$ . Nevertheless, the class  $V_\alpha$  is not exactly the set of orderings of rank  $\alpha$ . For instance, the ordering  $\omega + \omega$  is of rank 1 and belongs to  $V_2$ . Therefore, we work on slightly different inductive classes. For any  $\alpha \in \mathcal{O}$ , we define the class  $W_\alpha$  by :

$$W_\alpha = \left\{ \sum_{j \in J} K_j \mid J \in \mathcal{N} \text{ and } K_j \in V_\alpha \right\}.$$

Those classes are strictly intermediate to the Hausdorff's ones: the inclusions  $V_\alpha \subset W_\alpha \subset V_{\alpha+1}$  hold for any ordinal  $\alpha$ . For instance, the ordering  $\omega^\alpha + \omega^\alpha$  belongs to  $W_\alpha$  but does not belong to  $V_\alpha$  and the ordering  $\omega^{\alpha+1}$  belongs to  $V_{\alpha+1}$  but does not belong to  $W_\alpha$ . Formally, the *rank* of a linear ordering  $J$  is the smallest ordinal  $\alpha$  such that  $J \in W_\alpha$ . For instance the orderings of rank 0 are the finite ones. In this paper, we restrict to linear orderings of finite ranks that is the set  $\bigcup_{n < \omega} W_n = \bigcup_{n < \omega} V_n$ .

By extension, the rank of a word is the rank of its length and the rank of a set of words is the upper bound of the ranks of its elements.

We denote by  $A^\circ$  the set of all words indexed by countable scattered linear orderings and we also denote by  $A^{W_r}$  (respectively  $A^{V_r}$ ) the set of words whose length is an ordering in  $W_r$  (respectively  $V_r$ ) for some integer  $r$ . Thus the words of  $A^{W_r}$  have a rank lower than or equal to  $r$ .

### 3 Rational sets of words on linear orderings

Bruyère and Carton [2] have introduced rational expressions and automata for words indexed by countable scattered linear orderings. They have proved that a set of words is rational if and only if it is recognizable extending Kleene's theorem. More precisely, they have defined a whole hierarchy of rational sets [3]. For each subset of rational operations, they consider the class of corresponding rational languages and define transitions of automata capturing the same languages. In the following section, the characterization of rational sets of words of finite rank is notified.

#### 3.1 Rational expressions

The rational sets of finite rank can be obtained from finite sets of finite words using the union  $+$ , the concatenation  $\cdot$ , the star  $*$ , the omega iteration  $\omega$  and the backwards omega iteration  $-\omega$ . Let  $X$  and  $Y$  be two sets of words, we define:

$$\begin{aligned}
 X + Y &= \{z \mid z \in X \cup Y\} \\
 X \cdot Y &= \{x \cdot y \mid x \in X, y \in Y\} \\
 X^* &= \left\{ \prod_{j=1}^n x_j \mid n \in \mathcal{N}, x_j \in X \right\} \\
 X^\omega &= \left\{ \prod_{j \in \omega} x_j \mid x_j \in X \right\} \\
 X^{-\omega} &= \left\{ \prod_{j \in -\omega} x_j \mid x_j \in X \right\}
 \end{aligned}$$

To define rational sets of words indexed by all linear orderings, three more operations are needed : the ordinal iteration #, the backwards ordinal iteration -# and the iteration for all linear countable scattered orderings  $\diamond$ .

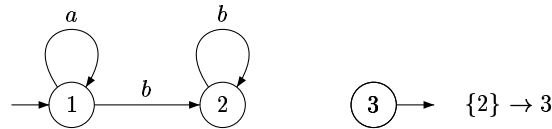
$$\begin{aligned}
 X^\# &= \left\{ \prod_{j \in J} x_j \mid J \in \mathcal{O}, x_j \in X \right\} \\
 X^{-\#} &= \left\{ \prod_{j \in -J} x_j \mid J \in \mathcal{O}, x_j \in X \right\} \\
 X \diamond Y &= \left\{ \prod_{j \in J \cup \hat{J}^*} z_j \mid J \in \mathcal{S} \setminus \emptyset, z_j \in X \text{ if } j \in J \text{ and } z_j \in Y \text{ if } j \in \hat{J}^* \right\}
 \end{aligned}$$

In this paper, we are only interested in languages which are defined using +, ·,  $\omega$  and  $-\omega$ . We refer the reader to [2] for a precise definition of other rational operations. A set of words on linear orderings is *rational* if it is obtained from finite sets of finite words using the rational operations defined above.

### 3.2 Automata on linear orderings

Let  $(Q, A, E, I, F)$  be a classical automaton on finite words with usual notations. As the set  $E$  of transitions is a subset of  $Q \times A \times Q$ , the paths of such an automaton are finite. In Büchi automata, a word is accepted if it is the label of a path going infinitely times through a given set of states. The problem is that this accepting condition does not even allow to recognize the concatenation of infinite words. To cope with this difficulty, a set of limit transitions included in  $\mathcal{P}(Q) \times Q$  is introduced. This way, if an infinite path goes infinitely many times through the states of a set  $P$  and that the transition  $(P, q)$  exists, then the next state of the path may be  $q$ .

*Example 1.* : Let  $\mathcal{A} = (Q, A, E, I, F)$  be the automaton of Figure 1 where  $Q = \{1, 2, 3\}$ ,  $A = \{a, b\}$ ,  $I = \{1\}$ ,  $F = \{3\}$ .



**Fig. 1.** Automaton recognizing  $a^*b^\omega$

A limit transition  $\{2\} \rightarrow 3$  is added to  $E$ . Intuitively, an infinite path going through the state 2 infinitely many times leads to state 3 and a path in  $\mathcal{A}$  leading from state 2 to state 3 is labelled  $b^\omega$ . Finally, this automaton recognizes the language  $a^*b^\omega$ .

The previous limit transitions called *left limit transitions* allow to recognize sets of words indexed by countable ordinals. In order to get words indexed by linear scattered orderings, we also need *right limit transitions*.

**Definition 1.** An automaton  $\mathcal{A}$  on linear orderings is defined by  $\mathcal{A} = (Q, A, E, I, F)$  where  $Q$  is a finite set of states,  $A$  is a finite alphabet,  $E \subseteq (Q \times A \times Q) \cup (\mathcal{P}(Q) \times Q) \cup (Q \times \mathcal{P}(Q))$  is the set of transitions and  $I \subseteq Q$  and  $F \subseteq Q$  are respectively the sets of initial and final states.

Right limit transitions are used symmetrically when a path has a limit length on the left. In order to use nested limit transitions, it is needed to define the left (respectively right) limit sets of states in a given point of the path.

Consider a finite path  $q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} q_n$  labelled  $x = a_1 \dots a_n$ . Note that a state is inserted between any two consecutive letters of  $x$ . In other words, to any two-factorization  $x = (a_1 \dots a_k)(a_{k+1} \dots a_n)$  of  $x$  is associated a state  $q_k$ . This definition of paths is generalized to automata on linear orderings in the following way: Let  $x$  be a word indexed by a linear scattered ordering  $J$ . To any two-factorization  $x = yz$  of  $x$ , one can associate a partition of  $J$  into two intervals  $(K, L)$  such that  $|y| = K$  and  $|z| = L$ . Then, a path labelled  $x$  is a function from the set  $\hat{J} = \{(K, L) \mid K \cup L = J \wedge \forall k \in K, \forall l \in L, k < l\}$  into the set of states. As the set  $\hat{J}$  is naturally equipped with the ordering  $(K_1, L_1) < (K_2, L_2)$  if and only if  $K_1 \subset K_2$ , a path labelled by a word of length  $J$  is a word over  $Q$  of length  $\hat{J}$ . An element of  $\hat{J}$  is called a *cut*.

Let  $\gamma = (q_c)_{c \in \hat{J}}$  be a word of length  $\hat{J}$  over  $Q$ , we are now able to define the limit sets of states of  $\gamma$  in a given cut  $c$  of  $\hat{J}$ :

$$\lim_{c^-} \gamma = \{q \in Q \mid \forall c' < c, \exists c'' < c' \text{ such that } q = q_{c''}\}$$

$$\lim_{c^+} \gamma = \{q \in Q \mid \forall c' > c, \exists c < c'' < c' \text{ such that } q = q_{c''}\}$$

For instance, in example 1, the word  $\gamma = (q_c)_{c \in \hat{\omega}}$  defined by  $q_{(\emptyset, \omega)} = 1$ ,  $q_{(\{0, 1, \dots, n\}, \{n+1, \dots\})} = 2$  for any positive integer  $n$  and  $q_{(\omega, \emptyset)} = 3$  has the following nonempty limit  $\lim_{(\omega, \emptyset)^-} \gamma = \{2\}$ .

Finally, a path has to be compatible with the automata transitions:

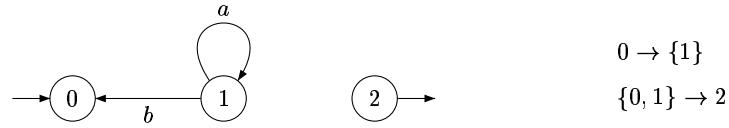
**Definition 2.** Let  $\mathcal{A} = (Q, A, E, I, F)$  be an automaton on linear orderings and let  $x = (a_j)_{j \in J}$  be a word of length  $J$  on  $A$ .

A path  $\gamma$  of label  $x$  in  $\mathcal{A}$  is a word  $\gamma = (q_c)_{c \in \hat{J}}$  of length  $\hat{J}$  over  $Q$  such that for any  $(K, L) \in \hat{J}$ :

- If there exists  $l \in L$  such that  $(K \cup \{l\}, L \setminus \{l\}) \in \hat{J}$   
then  $q_{(K, L)} \xrightarrow{a_l} q_{(K \cup \{l\}, L \setminus \{l\})} \in E$  else  $q_{(K, L)} \rightarrow \lim_{(K, L)^-} \gamma \in E$ .
- If there exists  $k \in K$  such that  $(K \setminus \{k\}, L \cup \{k\}) \in \hat{J}$   
then  $q_{(K \setminus \{k\}, L \cup \{k\})} \xrightarrow{a_k} q_{(K, L)} \in E$  else  $\lim_{(K, L)^+} \gamma \rightarrow q_{(K, L)} \in E$ .

Thus, if a cut has a predecessor or a successor, usual transitions are used, else the path is built on limit transitions. As  $\hat{J}$  has the least element  $(\emptyset, J)$  and the greatest element  $(J, \emptyset)$  for any linear ordering  $J$ , a path has always a first and a last state. It is said to be *successful* if it leads from an initial state to a final state. A word is *recognized* by an automata if it is the label of a successful path.

We denote by  $p \xrightarrow{x} q$  the existence of a path leading from state  $p$  to  $q$  of label  $x$ . The *content* of a path is the set of states occuring in the path and  $p \xrightarrow[x_P]{x} q$  denotes a path leading from  $p$  to  $q$  of label  $x$  and of content  $P$ .



**Fig. 2.** Automaton on linear orderings recognizing  $(a^{-\omega}b)^{\omega}$

### 3.3 Generalisations of Kleene’s theorem

Bruyère and Carton have generalized Kleene’s theorem on words indexed by countable scattered linear orderings:

**Theorem 2.** [2] *A set of words indexed by countable scattered linear orderings is rational if and only if it is recognizable.*

Moreover, they have defined a subclass of automata on linear orderings which recognizes rational languages of finite ranks.

**Theorem 3.** [3] *A set of words of finite rank is rational if and only if it is recognized by an automata on linear orderings where limit transitions  $P \rightarrow q$  or  $q \rightarrow P$  verify  $q \notin P$ .*

## 4 Complement of a rational set of finite rank

In the case of finite words, it is known that rational sets are closed under complementation. Given an automaton on finite words recognizing a language  $L$ , the construction of an automaton recognizing the complement  $A^* \setminus L$  is based on the property that any finite automaton on finite words can be determinized. Büchi has generalized this result for sets of words indexed by countable ordinals of finite ranks [9]. This property does not hold any longer for automata on linear orderings. An automaton on linear orderings  $\mathcal{A} = (Q, A, E, I, F)$  is *deterministic* if for any state  $q \in Q$  and any word  $u \in A^{\circ}$ , there exists at most one path labelled  $u$  starting from  $q$ .

**Proposition 1.** *The language  $(a^{-\omega})^{-\omega}$  can not be recognized by a deterministic automaton.*

To cope with this difficulty of determinism, we use a different method based on equivalence classes to prove the closure of rational sets under complementation. Up to now, we are only able to prove this result in the case of rational sets of words of finite ranks.

**Theorem 4.** *Let  $L$  be a rational set of words on linear orderings and let  $r$  be a finite integer. The complement  $A^{W_r} \setminus L$  is rational.*

In the case of finite words, Büchi has given a different proof of the closure under complement of rational sets. It does not need the property of determinizability but it is based on the following equivalence relation defined for any finite automaton  $\mathcal{A} = (Q, A, E, I, F)$  on finite words:

$$u \sim v \text{ if and only if } \forall p \in Q, \forall q \in Q, p \xrightarrow{u} q \iff p \xrightarrow{v} q$$

Note that if a word  $u$  is the label of a successful path in  $\mathcal{A}$ , it holds for any equivalent word. So any equivalence class is either contained in the language  $L$  recognized by  $\mathcal{A}$  or disjoint from  $L$ . Moreover, equivalence classes are rational thus the complement of  $L$  is rational as a finite union of equivalence classes. We extend this proof to automata on linear orderings of finite ranks. Let  $\mathcal{A} = (Q, A, E, I, F)$  be an automaton on linear orderings recognizing  $L$ . Recall that a path from  $p$  to  $q$  with label  $u$  and content  $P$  is denoted by  $p \xrightarrow[u]{P} q$ . As the contents of paths are needed in limit transitions, we define the equivalence relation  $\sim$  by:

$$u \sim v \text{ if and only if } \forall p \in Q, \forall q \in Q, \forall P \subseteq Q, p \xrightarrow[u]{P} q \iff p \xrightarrow[v]{P} q$$

Note first that the equivalence relation has finitely many classes. Indeed the class of a word  $u$  depends on whether there is a path from  $p$  to  $q$  with content  $P$  for each triple  $(p, q, P)$ . Since there are  $n^2 2^n$  such triples, the relation  $\sim$  has at most  $2^{n^2 2^n}$  equivalence classes. We denote by  $\mathcal{C}$  the set of all equivalence classes of  $\sim$ . For each integer  $r$ , we denote by  $\mathcal{C}_r = \{C \cap A^{W_r} \mid C \in \mathcal{C}\}$  the set of equivalence classes of rank  $r$ . The cardinality of  $\mathcal{C}_r$  is at most the cardinality of  $\mathcal{C}$ . As in the case of finite words, each class  $C$  is either contained in  $L$  or disjoint from  $L$ . Therefore we have both equalities

$$L = \bigcup_{C \in \mathcal{C}, C \cap L \neq \emptyset} C \quad \text{and} \quad \bar{L} = A^\circ \setminus L = \bigcup_{C \in \mathcal{C}, C \cap L = \emptyset} C$$

The same holds for words of rank less than  $r$ .

$$L \cap A^{W_r} = \bigcup_{C \in \mathcal{C}_r, C \cap L \neq \emptyset} C \quad \text{and} \quad A^{W_r} \setminus L = \bigcup_{C \in \mathcal{C}_r, C \cap L = \emptyset} C.$$

For each integer  $r$ , the family  $\mathcal{C}_r$  contains finitely many classes. To prove that  $A^{W_r} \setminus L$  is rational, it suffices to prove that each  $C \in \mathcal{C}_r$  is rational. We prove that claim by induction on  $r$ . The result holds obviously for  $r = 0$  and the induction step is based on the following idea. Suppose that  $\mathcal{C}_r$  contains the

classes  $\{C_1, \dots, C_m\}$ . We define rational expressions using the  $C_i$  as letters. An elementary expression is an expression of the form  $C_i$ ,  $C_i^\omega$  or  $C_i^{-\omega}$  where  $C_i$  is a class of  $\mathcal{C}_r$ . We denote by  $B$  the set of elementary expressions. We consider the set  $B^*$  of all expressions obtained by concatenation of elementary expressions. Suppose for instance that  $\mathcal{C}_r = \{C_1, C_2\}$ . The set of elementary expressions is  $B = \{C_1, C_1^\omega, C_1^{-\omega}, C_2, C_2^\omega, C_2^{-\omega}\}$  and a typical example of element of  $B^*$  is  $C_2^\omega C_1 C_2^{-\omega} C_1 C_2^{-\omega}$ . We consider each element of  $B^*$  as a rational expression over the letters  $C_i$ . Each expression of  $B^*$  denotes a set of words of rank at most  $r+1$ . By a slight abuse of language, we say that a word belongs to an expression  $R$  in  $B^*$  if it actually belongs to the set denoted by  $R$ . The two following lemmas are needed in the proof of proposition 2. Their proofs are not detailed in this paper because of the lack of space. In Lemma 1, we first prove that each word of rank at most  $r+1$  belongs to at least one expression in  $B^*$ .

**Lemma 1.**  $A^{W_{r+1}} = \bigcup_{R \in B^*} R$ .

In Lemma 2, we prove that two words belonging to the same expression are  $\sim$ -equivalent. This means that each set denoted by an expression of  $B^*$  is included in a single  $\sim$ -class.

**Lemma 2.** *If two words  $x, y$  of rank at most  $r+1$  belong to the same expression  $R$  of  $B^*$ , then they satisfy  $x \sim y$ .*

It follows from Lemmas 1 and 2 that each class  $C$  in  $\mathcal{C}_{r+1}$  satisfies

$$C = \bigcup_{R \in B^*, C \cap R \neq \emptyset} R$$

However, this is not a rational expression since there are infinitely many such expressions  $R$  included in  $C$ . In the following proposition, we show that the set of rational expressions included in some class  $C$  can be described by a rational expression over the elementary expressions.

**Proposition 2.** *Each equivalence class in  $\mathcal{C}_r$  is rational.*

The proof by induction on the rank  $r$  is not detailed in this paper. We come back to the proof of Theorem 4.

*Proof.* Let  $\mathcal{A}$  be an automaton on linear orderings recognizing  $L$  and let  $r$  be a finite rank. Let  $\mathcal{C}_r$  be the set of equivalence classes of rank  $r$  according to  $\mathcal{A}$ . From proposition 2, we have that each class of  $\mathcal{C}_r$  is rational. Moreover, considering the definition of  $\sim$ , we note that if a word  $u$  is the label of a successful path in  $\mathcal{A}$ , it holds for any equivalent word. So an equivalence class is either contained in  $L$  or disjoint of  $L$ . We deduce a rational expression of  $A^{W_r} \setminus L$  as a finite union of classes of  $\mathcal{C}_r$ :

$$A^{W_r} \setminus L = \bigcup_{C \in \mathcal{C}_r, C \cap L = \emptyset} C$$

□

As a conclusion, we mention a question that is left open by this paper. A generalization of our result is that the class of rational sets of countable scattered linear orderings is closed under complementation.

## References

1. A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *The Design and Analysis of Computer Algorithms*. Addison-Wesley, London, 1974.
2. V. Bruyère and O. Carton. Automata on linear orderings. In J. Sgall, A. Pultr, and P. Kolman, editors, *MFCS'2001*, volume 2136 of *Lect. Notes in Comput. Sci.*, pages 236–247, 2001. IGM report 2001-12.
3. V. Bruyère and O. Carton. Hierarchy among automata on linear orderings. In *IFIP TCS'2002*, pages 107–118, 2002.
4. N. Bedon. Finite automata and ordinals. *Theoret. Comput. Sci.*, 156:119–144, 1996.
5. N. Bedon. Automata, semigroups and recognizability of words on ordinals. *Int. J. Alg. Comput.*, 8:1–21, 1998.
6. N. Bedon and O. Carton. An Eilenberg theorem for words on countable ordinals. In Cláudio L. Lucchesi and Arnaldo V. Moura, editors, *Latin'98: Theoretical Informatics*, volume 1380 of *Lect. Notes in Comput. Sci.*, pages 53–64. Springer-Verlag, 1998.
7. J. R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik und grundl. Math.*, 6:66–92, 1960.
8. J. R. Büchi. On a decision method in the restricted second-order arithmetic. In *Proc. Int. Congress Logic, Methodology and Philosophy of science, Berkeley 1960*, pages 1–11. Stanford University Press, 1962.
9. J. R. Büchi. Transfinite automata recursions and weak second order theory of ordinals. In *Proc. Int. Congress Logic, Methodology, and Philosophy of Science, Jerusalem 1964*, pages 2–23. North Holland, 1965.
10. Y. Choueka. Finite automata, definable sets, and regular expressions over  $\omega^n$ -tapes. *J. Comput. System Sci.*, 17(1):81–97, 1978.
11. D. Girault-Beauquier. Bilimites de langages reconnaissables. *Theoret. Comput. Sci.*, 33(2–3):335–342, 1984.
12. F. Hausdorff. *Set theory*. In *Chelsea*, New York, 1957.
13. S. C. Kleene. Representation of events in nerve nets and finite automata. In C.E. Shannon, editor, *Automata studies*, pages 3–41. Princeton university Press, Princeton, 1956.
14. D. Muller. Infinite sequences and finite machines. In Proc. of Fourth Annual IEEE Symp., editor, *Switching Theory and Logical Design*, pages 3–16, 1963.
15. M. Nivat and D. Perrin. Ensembles reconnaissables de mots bi-infinis. In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*, pages 47–59, 1982.
16. D. Perrin and J.E. Pin. *Infinite words*. In Elsevier, editor, *Academic Press*, 2003.
17. J.-E. Pin. *Handbook of formal languages*, volume 1, chapter Syntactic semigroups, pages 679–746. Springer-Verlag, 1997.
18. M. O. Rabin. Decidability of second-order theories and automata on infinite trees. 141:1–35, 1969.
19. F. D. Ramsey. On a problem of formal logic. *Proc. of the London math. soc.*, 30:338–384, 1929.
20. J. G. Rosenstein. *Linear ordering*. Academic Press, New York, 1982.
21. S. Safra. On the complexity of  $\omega$ -automata. In *29th Annual Symposium on Foundations of computer sciences*, pages 24–29, 1988.
22. J. Wojciechowski. Finite automata on transfinite sequences and regular expressions. *Fundamenta informaticæ*, 8(3-4):379–396, 1985.