Theoretical Computer Science 108 (1993) 45–82
Elsevier

Synchronized rational relations of finite and infinite words*

Christiane Frougny

Université Paris 8 and Laboratoire Informatique Théorique et Programmation, Institut Blaise Pascal, 4 place Jussieu, 75252 Paris Cedex 05, France

Jacques Sakarovitch

CNRS, Institut Blaise Pascal, 4 place Jussieu, 75252 Paris Cedex 05, France

Abstract

Frougny, Ch. and J. Sakarovitch, Synchronized rational relations of finite and infinite words, Theoretical Computer Science 108 (1993) 45–82.

The purpose of this paper is a comprehensive study of a family of rational relations, both of finite and infinite words, namely those that are computable by automata where the reading heads move simultaneously on the n input tapes, and that we thus propose to call *synchronized rational relations*.

Introduction

Rational relations on words are those relations that are computable by finite automata with two or more 1-way input tapes. These relations have been considered by Rabin and Scott [20], who stated in their survey paper a number of results and problems, the last of which – the equivalence of deterministic n-tape automata – has been solved only recently [14]. The first systematic study is due to Elgot and Mezei [7], who established in their seminal paper of 1965 a number of properties of these relations, among which the closure by composition. Both these papers deal with finite words, whereas Büchi considered, for his investigation of second-order logic, computations of finite automata on n-tuples of infinite words [4].

The purpose of this paper is a comprehensive study of a family of rational relations, both of finite and infinite words, namely those that are computable by automata with

Correspondence to: Ch. Frougny, Institut Blaise Pascal, 4 place Jussieu, 75252 Paris Cedex 05, France.

* This work has been partly supported by the PRC Mathématiques et Informatique of the Ministère de la Recherche et de la Technologie and by the ESPRIT Basic Research Action 3166 ASMICS.

the main property that the reading heads move simultaneously on the n input tapes, and that we thus propose to call *synchronized rational relations*.

One of the main ideas developed in the paper is the *resynchronization* of automata with bounded delay – which holds for infinite computations as well as for the finite ones – and which can be stated as follows, using the Rabin and Scott model of n-tape automata. Let us say that an automaton has *bounded delay* if in the course of any computation the distance between the reading heads keeps bounded, and that an automaton is *synchronous* (or *letter-to-letter*) if the heads move simultaneously at every step of any computation. If an automaton $\mathscr{A}$ has bounded delay then it is possible to resynchronize it, i.e. to (effectively) build a synchronous automaton $\mathscr{A}'$ which is equivalent to $\mathscr{A}$ – for both finite and infinite computations.

Another side of the results presented in the paper – close to the idea of resynchronization but relevant only to finite words – is the link that can be established between some properties of the *result* of computations of an automaton and the properties of the *computations themselves* or, equivalently, between properties of a rational relation as a subset and properties of the rational (or regular) expression that realize it. This relationship corresponds exactly to the *Fatou property* in ring theory; it has been considered several times in the theory of rational relations (see [3, 6, 18] for instance). The best-known result along this line is probably the theorem first set out in [8] and usually credited to Eilenberg and Schützenberger (see [6, p. 265]), which states that a *length-preserving rational relation of* $A^* \times B^*$ is indeed a rational subset of $(A \times B)^*$ (cf. Theorem 2.2) or, equivalently, is realized by a synchronous automaton.

In this respect our paper deals not only with length-preserving relations – i.e. two related words have equal length – but, more generally, with relations with *bounded length difference* – i.e. the difference of the length of two related words in bounded. Rational relations with bounded length difference are characterized as a finite union of the products of length-preserving rational relations by finite sets (Proposition 2.1) and have the property that any n-tape automaton that realizes such a relation has bounded delay (Proposition 3.2), that brings back to resynchronization property.

Our first motivation for that work was the problem of representations of the integers (by finite words) and of the real numbers (by infinite words) in nonclassical numeration systems. The connection of this subject with the matter developed here is presented in Appendix B. In this context, the characteristic property of the relations under investigation – two words are in relation if they represent the same number – is that the difference of the length of two related words is bounded (when it comes to integers and to finite words), and that any automaton that realizes the relation has bounded delay (when it comes to reals and infinite words).

As far as the infinite words are concerned, the automata with bounded delay and the relations they realize, proved to be the "good" objects. But as for finite words, we were led to a natural generalization of relations with bounded delay, which happens to have been considered several times in the literature. We define *synchronized rational relations* (of finite words) to be the relations realized by letter-to-letter automata with terminal function in the family of rational sets (cf. Definition 4.1). The synchronized rational relations coincide with "finite-automata-definable relations" of Elgot and Mezei [7] and with "sequential" relations of Büchi [4] (see also [23]). Both of these old denominations are now in conflict with the (rather well accepted) terminology for rational relations (see [6, 2]).

As they make up an *effective Boolean algebra* (both in the case of finite and infinite words), synchronized rational relations (and its subfamily of relations with bounded length difference) appear in several fields related to automata theory. The multiplication by a generator in an automatic group is a rational relation with bounded length difference [8], the relations that describe ground rewriting systems are synchronized rational relations [5] and rational relations with bounded delay, both of finite and infinite words, have been studied in connection with logical hierarchies [23, 24], to mention a few.

To tell the truth – and as we discovered afterwards – most of the results on synchronized rational relations (of finite words) presented here can be found in [7]. The proofs we are giving are based on a key lemma of Eilenberg and Schützenberger, which we call the resynchronization lemma (Lemma 2.3) and generalize for infinite words (Lemma 7.9). These two lemmas deal with rational expressions; a direct algorithmic proof of the resynchronization of automata is presented in Appendix A. The main tool of our presentation is the notion of automata with terminal function. We also sketch a theory of letter-to-letter automata with terminal function. A complete description of the properties that the resynchronization lemma entails will be developed in a forthcoming paper.

The paper continues with the description of the two families of rational relations, of finite and infinite words, stressing on the situation of synchronized rational relations, in particular with respect to the deterministic ones.

For finite words, synchronized rational relations contain the recognizable relations and are contained in the deterministic rational relations. We show that it is not decidable whether a rational relation is synchronized or not (Proposition 5.5).

For infinite words, the situation is more intricate, for a nondeterministic automaton is not necessarily equivalent to a deterministic one. Using topological arguments, we show that the family of deterministic synchronized rational relations is the intersection of the family of synchronized rational relations and the family of deterministic rational relations (Theorem 8.7). The proof is based on the fact that the infinite behavior of a deterministic n-tape automaton is a countable intersection of open sets (Proposition 8.9) and it is noteworthy that the converse of this proposition does not hold, contrary to what happens for deterministic 1-tape automaton.

The undecidability results in the description of the family of rational relations of finite words carry over on the family of rational relations of infinite words, by means of the notion of *ω-completion* of a relation of finite words (Definition 8.13). The proof of results concerning deterministic relations, e.g. the fact that determinism is preserved by ω-completion, requires a number of direct constructions on automata, that are gathered in Appendix A. They are all simplified by the definition of *covering* of automata that is adapted from Stallings [22].

A preliminary version of this work has appeared in the *Proceedings of the International Colloquium on Words, Languages and Combinatorics* held in Kyoto, 1990, and published by World Scientific, and in [12].

PART I: RELATIONS OF FINITE WORDS

1. Preliminaries

We basically follow the exposition of [7] and of [6] for the definition of finite automata over an alphabet and their generalization to automata over any monoid M; when M is a direct product of (n) free monoids – which will be the only case considered here – this generalization amounts to the definition of n-tape automata.

An *automaton over a finite alphabet* A, $\mathscr{A}=(Q, A, E, I, T)$ is a directed graph labelled by elements of A; Q is the set of *states*, $I\subset Q$ is the set of *initial* states, $T\subset Q$ is the set of *terminal* states and $E\subset Q\times A\times Q$ is the set of labelled *edges*. The automaton is *finite* if Q is finite and this will always be the case in this paper. If $(p, a, q)\in E$, we note also $p\xrightarrow{a}q$. A *computation* c in $\mathscr{A}$ is a finite sequence of labelled edges

$$c=p_0\xrightarrow{a_1}p_1\xrightarrow{a_2}p_2\cdots\xrightarrow{a_n}p_n.$$

The *label* of c, denoted by $|c|$, is the element $|c|=a_1a_2\ldots a_n$ of A^*. The computation c is *successful* if $p_0\in I$ and $p_n\in T$. The *behavior* of $\mathscr{A}$ is the subset $|\mathscr{A}|$ of A^* consisting of labels of successful computations of $\mathscr{A}$.

An automaton $\mathscr{A}$ is said to be *trimmed* if (a) every state q is *accessible*, i.e. there exists a path in $\mathscr{A}$ starting in I and terminating in q and (b) every state q is *coaccessible*, i.e. there exists a path in $\mathscr{A}$ starting in q and terminating in T.

This definition of automata as labelled graphs extends readily to automata over any monoid: an *automaton over* M, $\mathscr{A}=(Q, M, E, I, T)$, is a directed graph the edges of which are labelled by elements of the monoid M. The automaton is finite if the set of edges $E\subset Q\times M\times Q$ is finite (and, thus, Q is finite). The label of a computation

$$c=p_0\xrightarrow{x_1}p_1\xrightarrow{x_2}p_2\cdots\xrightarrow{x_n}p_n$$

is the element $|c|=x_1x_2\ldots x_n$ of M. The *behavior* of $\mathscr{A}$ is the subset $|\mathscr{A}|$ of M consisting of labels of successful computations of $\mathscr{A}$. In this context, an automaton over an alphabet A is an automaton over the free monoid A^*.

The set of behaviors of finite automata over a monoid M coincides with the Kleene closure of finite sets of M – this is Kleene's theorem for M being a free monoid and can be found in [7] for the general case. Let us state this in the vocabulary we shall be using. Let M be a monoid; the family of *rational* subsets of M is the least family of subsets of M containing the finite subsets and closed under union, product and the "star" operation (and is denoted by $\operatorname{Rat} M$). Then the following theorem holds.

Theorem 1.1 (Elgot and Mezei [7]). *A subset of M is rational if and only if it is the behavior of a finite automaton over M, the labels of the edges of the automaton being taken in any set of generators of M.*

We shall consider only finite automata and, thus, call them simply *automata* in the sequel. By definition, a *rational relation* from a free monoid A^* into a free monoid B^* is a relation the graph of which is a rational subset of (the nonfree monoid) $A^* \times B^*$. The family of rational relations from A^* into B^* is, thus, denoted by $\operatorname{Rat}(A^* \times B^*)$, and by Rat_2 when the alphabets are not specified. By Theorem 1.1, a rational relation is the behavior of a (finite) automaton over $A^* \times B^*$. The labels of its edges are pairs of words: such an automaton may be viewed as (or is equivalent to) a finite automaton with two tapes – the model defined by Rabin and Scott – and will be called here a 2-*automaton*, by reference to automata over the free monoid that are automata with one tape or 1-automata. In the literature 2-automata are also often called *nondeterministic generalized sequential machines* or *transducers*.

And now some more notations. The identity of a monoid is (ambiguously) denoted by 1. The length of a word $f \in A^*$ is denoted by $|f|$ and the number of occurrences of the letter a in f by $|f|_a$. The set of words of A^* of length k is denoted by A^k and the set of words of A^* of length $\leqslant k$ is denoted by $A^{\leqslant k}$. For every integer k, let us denote by Diff_k the set

$$\mathrm{Diff}_k = (A^{\leqslant k} \times 1) \cup (1 \times B^{\leqslant k}).$$

The set of natural integers is denoted by $\mathbf{N}$. Let $\lambda : A^* \times B^* \to \mathbf{N}^2$ be the length morphism $\lambda : (f, g) \mapsto (|f|, |g|)$.

Finally, let us make an observation of importance for the rest of the paper. The set of pairs of words in $A^* \times B^*$ of equal length is a *free submonoid* of $A^* \times B^*$ generated by $A \times B$.

Definition 1.2. We call *letter-to-letter 2-automaton* a 2-automaton with edges labelled in $A \times B$.

A letter-to-letter 2-automaton can, thus, be viewed as *a 1-automaton with input alphabet $A \times B$* and, in particular, can be determinized.

2. Rational relations with bounded length difference

We call *length difference* of a pair of words (f, g) the integer $||f| - |g||$. Any element $u = (f, g)$ of $A^* \times B^*$ with length difference $||f| - |g|| = k$ can be uniquely written as a product $u=u'u''$ with u' in the free monoid $(A\times B)^*$ and $u''\in\mathrm{Diff}_k$. This unique decomposition yields the following equality:

$$\forall T_1,T_2\subset(A\times B)^*,\ \forall S_1,S_2\subset\mathrm{Diff}_k,\quad T_1S_1\cap T_2S_2=(T_1\cap T_2)(S_1\cap S_2). \qquad (1)$$

A relation R has a *bounded length difference* if there exists an integer k such that the length difference of every pair (f,g) in R is smaller than or equal to k. It is also convenient to call *length difference* of a relation R the upper bound of the set of length differences of pairs of words in R. With that terminology, a relation has a bounded length difference if its length difference is finite and a relation is *length-preserving* if its length difference is 0.

In this section, we characterize bounded length difference rational relations. The results are basically due to Elgot and Mezei [7]. Our presentation, together with the proofs, rather follows the ideas of Eilenberg and Schützenberger and relies on a key lemma of [6] that is well-suited for the extension to relations of infinite words. Another possible proof is given by the direct resynchronization of automata (see Section A.2).

Proposition 2.1. *A rational relation R with length difference bounded by k is a finite union of products of rational subsets of $(A\times B)^*$ by subsets of* Diff_k *(a finite union of products of subsets of* Diff_k *by rational subsets of $(A\times B)^*$).*

The restriction of Proposition 2.1 to length-preserving relations gives the better-known statement.

Theorem 2.2 (Elgot and Mezei [7]). *A length-preserving rational relation of $A^*\times B^*$ is a rational subset of $(A\times B)^*$.*

Theorem 2.2 also appears in [6]. As a consequence of it (and of Theorem 1.1), a length-preserving rational relation is the behavior of a letter-to-letter 2-automaton (and this is closer to the original statement in [7]).

The proof of Proposition 2.1 uses the following lemma.

Lemma 2.3 (Eilenberg and Schützenberger [6], Resynchronization lemma). *For every $T\in\mathrm{Rat}(A\times B)^*$ and every word $w\in A^*$ there exists a family $\{T_x\mid x\in A^{|w|}\}$ of rational sets of $(A\times B)^*$ such that*

$$(w,1)T=\bigcup_{x\in A^w}T_x(x,1).$$

Proof of Proposition 2.1. It mimics the proof of Theorem 2.2 as given in [6] and goes by induction on the star height of R. If R is of star height 0, it is a finite union of pairs with length difference k, each one being the product of an element of $(A\times B)^*$ and of an element of Diff_k. Assume now that the proposition holds for every rational relation of star height h, and let R be of star height $h+1$. R is a finite union $R=\bigcup u_0 R_1^* u_1 \ldots u_{r-1} R_r^* u_r$, with the u_i's in $A^* \times B^*$ and the R_i's in $\mathrm{Rat}(A^* \times B^*)$ of star height at most h. Elementary length considerations show that every R_i has to be length-preserving and, thus, $R_i \in \mathrm{Rat}(A \times B)^*$ and, therefore, $R_i^* \in \mathrm{Rat}(A \times B)^*$. From Lemma 2.3, $u_0 R_1^*$ is a finite union of elements of the form Hv, with $H \in \mathrm{Rat}(A \times B)^*$ and $v \in A^* \times B^*$, and $\lambda(v)=\lambda(u_0)$.

Iterating this procedure for i going from 1 to r shows that the monomial $u_0 R_1^* \ldots u_{r-1} R_r^* u_r$ and, thus, R, is a finite union of elements of the form Lw, with $L \in \mathrm{Rat}(A \times B)^*$ and $\lambda(w)=\lambda(u_0 \ldots u_r)$. Thus, w has length difference k and may be written as $w=w'w''$, with $w' \in (A \times B)^*$ and $w'' \in \mathrm{Diff}_k$. □

Proposition 2.1 can equivalently be stated as follows:

"*A rational relation with bounded length difference is equal to a finite union of products of length-preserving rational relations by finite relations.*"

In order to express Proposition 2.1 in terms of automata, we define the notion of (2)-automaton with terminal (or initial) function; it makes use of the ideas introduced in [21] by Schützenberger for the definition of subsequential transducers (see also [2]).

Definition 2.4. A *2-automaton with terminal function* is a 2-automaton $\mathscr{A}=(Q, A^* \times B^*, E, I, \omega)$, where ω is a function from Q into $\mathscr{P}(A^* \times B^*)$. The *behavior of* $\mathscr{A}$, $|\mathscr{A}|$, is the set of pairs of words (f, g) such that there exist $f', f'' \in A^*$, $g', g'' \in B^*$ such that $f=f'f''$, $g=g'g''$, states p and q such that $p \in I$, $p \xrightarrow{(f',g')} q$ is a computation of $\mathscr{A}$ and $\omega(q)=(f'', g'')$.

A 2-automaton $\mathscr{A}=(Q, A^* \times B^*, E, \alpha, T)$ with *initial* function and a 2-automaton $\mathscr{A}=(Q, A^* \times B^*, E, \alpha, \omega)$ with *initial and terminal* function are defined similarly. It is clear that 2-automata with initial and/or terminal functions in $\mathrm{Rat}(A^* \times B^*)$ have the same power as classical 2-automata. (See Section 4 for the connection between terminal functions and the use of blank symbol at the end of the input tapes.)

Proposition 2.1 is then restated in terms of automata.

Corollary 2.5. *A rational relation with length difference bounded by k is the behavior of a letter-to-letter 2-automaton with a terminal function into subsets of* Diff_k *(resp. with an initial function into subsets of* Diff_k*).*

Fact 2.6 (*Determinization of letter-to-letter 2-automata with terminal function*). Let $\mathscr{A}=(Q, A \times B, E, I, \omega)$ be a letter-to-letter 2-automaton with a terminal function ω. Since $\mathscr{A}$ is indeed a 1-automaton on the alphabet $A \times B$, it can be determinized by the usual subset construction: $\mathscr{B}=(\mathscr{P}(Q), A \times B, F, \{I\}, \Omega)$. The terminal function Ω, defined by

$$\forall P \subset Q, \quad \Omega(P)=\bigcup_{q \in P} \omega(q),$$

makes $\mathscr{B}$ equivalent to $\mathscr{A}$.

Corollary 2.5 brings more than a reformulation of Proposition 2.1: together with (1) it yields the following corollary.

Corollary 2.7 (Elgot and Mezei [7]). *For every integer k, the family of rational relations with length difference smaller than k is (effectively) closed under intersection and set difference.*

Proof. Let R_i, $i=1,2$, be a rational relation realized by the letter-to-letter 2-automaton $\mathscr{A}_i=(Q_i, A\times B, E_i, I_i, \omega_i)$. Let $\mathscr{C}$ be the classical cartesian product of the two automata $\mathscr{A}_1$ and $\mathscr{A}_2$: $\mathscr{C}=(Q_1\times Q_2, A\times B, E, I_1\times I_2, \omega)$, where E is defined by $((p_1,p_2),(a,b),(q_1,q_2))\in E$ if and only if $(p_1,(a,b),q_1)\in E_1$ and $(p_2,(a,b),q_2)\in E_2$ and where ω is the "intersection" of the terminal functions ω_1 and ω_2:

$$\forall (q_1,q_2)\in Q_1\times Q_2,\quad \omega(q_1,q_2)=\omega_1(q_1)\cap\omega_2(q_2).$$

From (1) it follows easily that the behavior of $\mathscr{C}$ is exactly the intersection $R_1\cap R_2$.

If R is a rational relation with length difference smaller than k then $(A\times B)^*\mathrm{Diff}_k\backslash R$ is a rational relation (with length difference smaller than k). Let $\mathscr{A}=(Q, A\times B, E, I, \omega)$ be a letter-to-letter 2-automaton that realizes R. From what is said in Fact 2.6, $\mathscr{A}$ may be chosen *deterministic* and *complete*.

The automaton $\mathscr{A}'=(Q, A\times B, E, I, \omega')$ deduced from $\mathscr{A}$ by taking the "complementary" terminal function (with respect to Diff_k), i.e.

$$\forall q\in Q,\quad \omega'(q)=\mathrm{Diff}_k\backslash\omega(q)$$

will realize $(A\times B)^*\mathrm{Diff}_k\backslash R$. The closure under set difference follows since $R_1\backslash R_2= R_1\cap[(A\times B)^*\mathrm{Diff}_k\backslash R_2]$ if R_1 and R_2 have both length difference smaller than or equal to k. $\square$

3. 2-automata with bounded delay

Let R be a relation on $A^*\times B^*$. The pairs (f,g) in R are obtained by means of computations of a machine. Any statement on the length difference of a pair (f,g) is a statement on the *result* of one computation; any statement on the length difference of R is a statement on the set of results of all possible computations of that machine. The purpose of this section is to state properties of the *computations* themselves. Beyond the intrinsic interest, this change of point of view will be of importance in the sequel of the paper, on infinite words, when the corresponding statements on the result of the computation will no longer be possible.

A rational relation is realized by a finite automaton with two tapes and, thus, with two (reading) heads. Proposition 3.2 states that *during any computation of any 2-automaton that realizes a rational relation with bounded length difference, the distance between the two heads keeps bounded* (by an integer that obviously depends on the

2-automaton considered). The following definitions are taken in order to make this last sentence precise in the model of automata as labelled graphs.

Let $\mathcal{A}$ be a 2-automaton. Let c be a computation in $\mathcal{A}$:

$$c = q_0 \xrightarrow{(x_1, y_1)} q_1 \cdots q_{n-1} \xrightarrow{(x_n, y_n)} q_n \cdots q_{r-1} \xrightarrow{(x_r, y_r)} q_r.$$

We call *delay of* c *at step* n the integer

$$\mathrm{del}(c, n) = \big| |x_1 \ldots x_n| - |y_1 \ldots y_n| \big|;$$

that is, the absolute value of the difference of the length of the words read on the two tapes after the nth transition of the automaton. The *delay* of c is the maximum of the delays along the path

$$\mathrm{del}(c) = \max\{\mathrm{del}(c, n) \mid 1 \leqslant n \leqslant r\}.$$

The *delay* of a 2-automaton $\mathcal{A}$ is the supremum of the delays of all successful computations:

$$\mathrm{del}(\mathcal{A}) = \sup\{\mathrm{del}(c) \mid c \text{ successful computation in } \mathcal{A}\}.$$

In particular, a trimmed 2-automaton has a delay equal to 0 if and only if the length difference of the label of every edge is 0 – and is, thus, obviously equivalent to a letter-to-letter 2-automaton.

Lemma 3.1. *A trimmed 2-automaton has a bounded delay if and only if the length difference of the label of every loop is* 0.

Proof. Let $l = p \xrightarrow{(x, y)} p$ be a loop in $\mathcal{A}$ such that $|x| - |y| = k \neq 0$. Since $\mathcal{A}$ is trimmed, there exists a successful computation $c = q \xrightarrow{(u, v)} p \xrightarrow{(z, w)} t$ with $q \in I$ and $t \in T$. It is, thus, possible to build a sequence of successful computations

$$c_n = q \xrightarrow{(u, v)} p \xrightarrow{(x^n, y^n)} p \xrightarrow{(z, w)} t$$

with increasing delay.

Conversely, suppose that the label of every loop in $\mathcal{A}$ has length difference 0. Let N be the number of states of $\mathcal{A}$. Let

$$c = q_0 \xrightarrow{(x_1, y_1)} q_1 \cdots q_{n-1} \xrightarrow{(x_n, y_n)} q_n \cdots q_{r-1} \xrightarrow{(x_r, y_r)} q_r$$

be any computation in $\mathcal{A}$. Then $\mathrm{del}(c, n)$, viewed as a function from $\mathbf{N}$ into itself, can take at most N different values, for $\mathrm{del}(c, n)$ has the same value as the delay at the first occurrence of the state q_n in c. Moreover, this value can be computed on any subpath of c that goes from q_0 to q_n and that is a simple path. There are only a finite number of simple paths in $\mathcal{A}$, and this concludes the proof. $\square$

Now, if a relation R is realized by a (trimmed) 2-automaton with an unbounded delay, then R cannot have a bounded length difference. This, together with Lemma 3.1, yields the following.

Proposition 3.2. *A rational relation R has a bounded length difference if and only if any trimmed 2-automaton that realizes R has a bounded delay.*

Corollary 3.3. *The length difference of a rational relation is computable.*

Proof. A rational relation is given by a rational expression on $A^* \times B^*$ or, equivalently, by a 2-automaton. It is then sufficient to test whether every simple loop has a length-preserving label; and there is only a finite number of simple loops. If it is so, the length difference of the relation will be given by the maximum of the length difference of labels of simple paths from an initial state to a terminal state; otherwise, the length difference of the relation is infinite. $\square$

Note that the length difference of a rational relation can be computed as well on any rational expression that describes it.

The name given to Lemma 2.3 is made clear when the lemma is applied to automata. A 2-automaton $\mathcal{A}$ gives rise, by the classical (effective) procedure, to a rational expression; if $\mathcal{A}$ has a bounded delay, this expression describes a relation with bounded length difference and is transformed, by the resynchronization lemma, into a finite union of products of rational expressions over $A \times B$ by subsets of Diff_k and the latter is transformed again by (effective) classical procedure into a letter-to-letter 2-automaton with terminal function into Diff_k. Since the proof of the resynchronization lemma is effective (see [6]), the whole process is effective and we have Corollary 3.4.

Corollary 3.4. *Given a 2-automaton with bounded delay, an equivalent letter-to-letter 2-automaton with terminal (initial) function is computable.*

An "automatic" version of Lemma 2.3 yields an effective procedure for Corollary 3.4, without going back and forth through the equivalent rational expression (see Appendix A).

4. Synchronized rational relations

A natural generalization of rational relations with bounded length difference has already been considered in the literature several times.

Let $\mathrm{Diff}_{\mathrm{Rat}} = \{(S \times 1) \cup (1 \times T) \mid S \in \mathrm{Rat}\, A^*, T \in \mathrm{Rat}\, B^*\}$. That is to say, $\mathrm{Diff}_{\mathrm{Rat}}$ is the generalization of Diff_k from finite sets to rational sets.

Definition 4.1. A rational relation R is (left)-*synchronized* if it is realized by a letter-to-letter 2-automaton with a terminal function taking its value in $\mathrm{Diff}_{\mathrm{Rat}}$.

This definition can be rephrased as a rational relation is left-synchronized if and only if it is a finite union of products of length-preserving rational relations by elements of $\mathrm{Diff}_{\mathrm{Rat}}$.

It is very easy to verify that automata with terminal functions into $\mathrm{Diff}_{\mathrm{Rat}}$ have exactly the same behavior as 2-tape automata where the shorter input word is completed on the right side by enough blank symbols to match in length the longer one. The latter version of automata is used in [7] and in more logic-oriented papers like [4, 23, 24].

From Proposition 2.1 – and not from Definition 4.1 – Corollary 4.2 follows.

Corollary 4.2. *A rational relation with bounded length difference is a synchronized rational relation.*

Remark 4.3. Right-synchronized rational relations may be defined symmetrically. In contrast to what happens for rational relations with bounded length difference (which are at the same time having bounded right and left length differences), a right-synchronized rational relation is not necessarily left-synchronized: the relation $R=(a, a)^*(1, b)^*$, for instance, is a left-synchronized rational relation which is not right-synchronized.

Remark 4.4. Definition 4.1 holds for 2-ary relations. Whereas the definition of rational relations with bounded length difference readily extends from 2-ary to n-ary relations, the same generalization for synchronized rational relations, without being difficult, requires some attention. Let us denote by Synch_n the family of n-ary synchronized rational relations, with the convention that an n-ary relation can be considered as well as an m-ary relation (for $m \geqslant n$), with the remaining $m-n$ components equal to 1. For instance, Synch_1 is not only another notation for Rat but also for $\mathrm{Diff}_{\mathrm{Rat}}$ (in the case $m=2$), etc. With this convention, a rational n-relation is synchronized if it is realized by a letter-to-letter n-automaton with a terminal function taking its value in Synch_{n-1}.

Proposition 4.5 (Elgot and Mezei [7]). *The family of synchronized rational relations forms an effective Boolean algebra (in which emptiness is decidable).*

Proof. As noted in Fact 2.6, we can make the assumption that the relation R is realized by a deterministic (and complete) letter-to-letter 2-automaton $\mathscr{A}=(Q, A\times B, E, q_{\_}, \omega)$. The relation $\bar{R}$, complement of R, is then realized by the 2-automaton $\bar{\mathscr{A}}=(Q, A\times B, E, q_{\_}, \bar{\omega})$, where for every state q in Q, $\bar{\omega}(q)$ is the complement of $\omega(q)$ in $(A^*\times 1)\cup(1\times B^*)$. If $\omega(q)$ is in $\mathrm{Diff}_{\mathrm{Rat}}$, so is $\bar{\omega}(q)$ since Rat is closed under complementation. The same argument proves that Synch_n is an effective Boolean algebra, using an induction on n.

The legitimacy of these two constructions follows from the unique factorization property of the elements of $A^* \times B^*$ in the product $(A \times B)^*[(A^* \times 1) \cup (1 \times B^*)]$. $\square$

Fact 4.6 (*Minimization of deterministic letter-to-letter 2-automata with terminal function*). It seems to be noteworthy that, as the determinization, the minimization of automata extends from classical 1-automata to letter-to-letter 2-automata.

We sketch here the main steps of this generalization.

(a) Let R be any subset of $A^* \times B^*$. For every u in $(A \times B)^*$ we define the set of *continuations* of u in R:

$$\mathrm{Cont}_R(u) = \{w \in (A^* \times 1) \cup (1 \times B^*) \mid uw \in R\}.$$

Such an R defines an equivalence relation ρ_R on $(A \times B)^*$ by

$$u \simeq v \bmod \rho_R \Leftrightarrow \forall t \in (A \times B)^* \ \mathrm{Cont}_R(ut) = \mathrm{Cont}_R(vt).$$

(b) It is then a mere routine to check that ρ_R is a right regular equivalence of $(A \times B)^*$ and that ρ_R is coarser than the equivalence relation on $(A \times B)^*$ defined by any deterministic letter-to-letter 2-automaton with terminal function that recognizes R. The relation ρ_R itself defines an automaton $\mathscr{A}_R = (Q_R, A \times B, E_R, \{i\}, \omega_R)$, where Q_R is the quotient of $(A \times B)^*$ by ρ_R, $E_R = \{([u]_{\rho_R}, (a, b), [u(a, b)]_{\rho_R}) \mid [u]_{\rho_R} \in Q_R, (a, b) \in A \times B\}$ and $\omega_R([u]_{\rho_R}) = \mathrm{Cont}_R(u)$. $\mathscr{A}_R$ is the *minimal* deterministic letter-to-letter 2-automaton with terminal function of R.

(c) The minimal automaton $\mathscr{A}_R$ is effectively computable from any deterministic automaton $\mathscr{A} = (Q, A \times B, E, \{i\}, \omega)$ that recognizes R by a process that mimics the classical computation of the Nerode's equivalence. Let $\mathscr{P}_0, \mathscr{P}_1, \ldots$ be the sequence of partitions of Q defined by

- $p, q \in Q$, $p \simeq q \bmod \mathscr{P}_0$ if and only if $\omega(p) = \omega(q)$,
- $\mathscr{P}_{i+1}$ is a thinning of $\mathscr{P}_i$ and two states p and q are joined in $\mathscr{P}_i$ and separated in $\mathscr{P}_{i+1}$ if there exists (a, b) in $A \times B$ such that $\omega(p.(a, b)) \neq \omega(q.(a, b))$, where $p.u$ denotes the state r such that there is a path labelled by u from p to r. The length of the sequence is bounded by $|Q|$ and if $\mathscr{P}_{i+1} = \mathscr{P}_i$ then $\mathscr{P}_{i+n} = \mathscr{P}_i$ for any integer n. $\mathscr{A}_R$ is equal to the quotient of $\mathscr{A}$ by the thinnest partition $\mathscr{P}_i$.

Note that a letter-to-letter 2-automaton (with terminal function) realizes a relation that is both of finite image and of finite preimage if and only if the terminal function maps every state onto a pair of finite sets and, thus, realizes a relation with bounded length difference. Together with Proposition 2.1, this gives a simple proof of the following corollary.

Corollary 4.7 (Elgot and Mezei [7, Theorem 6.1]). *A rational relation has a bounded length difference if and only if it is synchronized and both of finite image and finite preimage.*

5. On the geography of Rat_2

We describe now the situation of the family of synchronized rational relations, denoted by Synch_2, within Rat_2. The undecidability result appears to be new. We recall first two definitions.

A *recognizable* relation of $A^* \times B^*$ is a finite union of cartesian products of the form $S \times T$ with S in $\mathrm{Rat}\, A^*$ and T in $\mathrm{Rat}\, B^*$ (see [2]). The set of these relations is denoted by $\mathrm{Rec}(A^* \times B^*)$ and simply by Rec_2 if the alphabets are not determined.

Let us now recall the definition of a *deterministic* 2-*automaton* (see [15] for instance). Let $\mathscr{A}$ be a 2-automaton on $A^* \times B^*$. It is always possible to suppose that the edges are labelled by elements of $(A \times 1) \cup (1 \times B)$. Up to now, the model of automaton we have considered works without endmarkers and is still as powerful as the other models in the literature. In the case of deterministic 2-automata, it is necessary – in order to make them powerful enough – to suppose that every word is followed on its tape by an endmarker, say \$, which belongs to neither A nor B, and that the reading of this endmarker induces a transition in the 2-automaton. A 2-automaton $\mathscr{A} = (Q, ((A \cup \$) \times 1) \cup (1 \times (B \cup \$)), E, \{i\}, T)$ is said to be *deterministic* if the following conditions hold:

(i) $Q = Q_A \cup Q_B$ is a partition such that

$$\forall q \in Q_A, \quad (q, e, p) \in E \;\Rightarrow\; e \in (A \cup \$) \times 1,$$

$$\forall q \in Q_B, \quad (q, e, p) \in E \;\Rightarrow\; e \in 1 \times (B \cup \$),$$

(ii) there is only one initial state $\{i\}$,

(iii) $\mathscr{A}$ is deterministic (in the usual meaning) on every tape.

The relation recognized by $\mathscr{A}$ on $A^* \times B^*$ is equal to $\{(u, v) \mid (u\$, v\$) \in |\mathscr{A}|\}$. A rational relation is said to be *deterministic* if it is recognized by a deterministic 2-automaton.

This definition makes our deterministic 2-automaton (defined as labelled graphs) equivalent to the deterministic 2-tape automata of Rabin and Scott [20]. In contrast with 1-automata, a 2-automaton needs not to be equivalent to a deterministic one. Deterministic rational relations, denoted by DRat_2, form a proper subclass of Rat_2, which is closed under complementation [20].

The following situation then holds (see also [15, 23]).

Theorem 5.1. $\mathrm{Rec}_2 \subset \mathrm{Synch}_2 \subset \mathrm{DRat}_2 \subset \mathrm{Rat}_2$.

The inclusions are strict: the identity relation belongs to $\mathrm{Synch}_2 \setminus \mathrm{Rec}_2$ and $R = (a, aa)^*$, for instance, to $\mathrm{DRat}_2 \setminus \mathrm{Synch}_2$.

The first two inclusions in Theorem 5.1 are also stated in [23]. The third one is well known (the endmarker may be replaced by a transition labelled by the empty word).

The second inclusion is established by the following proposition.

Proposition 5.2. *A synchronized rational relation is a deterministic rational relation.*

Proof. By definition, a synchronized rational relation is realized by a letter-to-letter 2-automaton (with terminal rational function). Every letter-to-letter 2-automaton is equivalent to a deterministic 1-automaton over $A \times B$. We, thus, have to show how to transform a deterministic letter-to-letter 2-automaton $\mathscr{A} = (Q, A \times B, E, \{i\}, \omega)$ into a deterministic 2-automaton $\mathscr{B}$. The automaton $\mathscr{B}$ is made of two parts: the first one realizes the transitions of $\mathscr{A}$, the second one takes into account the recognition of all rational sets that appear as image of the terminal function ω, while the transition to that part is induced by the reading of the endmarker on one tape or the other.

More precisely, let

$$\mathscr{B} = (Q_1 \cup Q_2 \cup \Sigma, ((A \cup \$) \times 1) \cup (1 \times (B \cup \$)), F \cup \Phi, \{i_B\}, T).$$

The first part of $\mathscr{B}$, $Q_1 \cup Q_2$ and F, is defined by $Q_2 = Q \times A$, $Q_1 = Q$ and for every edge $(p, (a, b), q)$ in E one has

$$(p_1, (a, 1), (p_2, a)) \in F \quad \text{and} \quad ((p_2, a), (1, b), q_1) \in F.$$

Let $q \in Q$ and $\omega(q) = (K_q \times 1) \cup (1 \times L_q)$. Let, first, $\mathscr{T}_q$ be a deterministic automaton (with initial state j_q) that recognizes L_q, and, for every a in A, denote by $\mathscr{S}_{a,q}$ a deterministic automaton (with initial state $l_{a,q}$) that recognizes $a^{-1}K_q$. Then Σ is the union of the states of the $\mathscr{T}_q$ and of the $\mathscr{S}_{a,q}$, and Φ is the set consisting of

(i) the union of the edges of the $\mathscr{T}_q$ with the labels read on the second tape [i.e. if (x, a, y) is an edge of $\mathscr{T}_q$ then $(x, (1, a), y)$ is in Φ] plus an edge $(q_1, (\$, 1), j_q)$ for every q in Q and

(ii) the union of the edges of the $\mathscr{S}_{a,q}$ with the labels read on the first tape [i.e. if (x, b, y) is an edge of $\mathscr{S}_{a,q}$ then $(x, (b, 1), y)$ is in Φ] plus an edge $((q_2, a), (1, \$), l_{a,q})$ for every a in A and for every q in Q.

The set T of terminal states of $\mathscr{B}$ is the union of the terminal states of the $\mathscr{T}_q$ and $\mathscr{S}_{a,q}$. It is easy to check that $\mathscr{B}$ is deterministic and equivalent to $\mathscr{A}$. □

The first inclusion in Theorem 5.1 is given by the following proposition.

Proposition 5.3. *Any recognizable relation is a (left- or right-) synchronized rational relation.*

Proof. Without loss of generality, we assume that $R = S \times T$ with S in Rat A^* and T in Rat B^*. S is recognized by the automaton $\mathscr{A} = (P, A, E, P_-, P_+)$ and T is recognized by the automaton $\mathscr{B} = (Q, B, F, Q_-, Q_+)$. For every state p in P let S_p be the set of words of A^* which are labels of a computation of $\mathscr{A}$ starting in p and ending in P_+, and let S'_p be the set of words of A^* which are labels of a computation of $\mathscr{A}$ starting in

P_- and ending in p. It holds that $S=\bigcup_{p\in P_-} S_p=\bigcup_{p\in P_+} S'_p=\bigcup_{p\in P} S'_p S_p$. For every q in Q, the sets T_q and T'_q are defined analogously. It then follows that

$$S\times T=\{w\in[(S'_p\times T'_q)\cap(A\times B)^*](1\times T_q)\mid p\in P_+,\ q\in Q\}$$

$$\cup\{w\in[(S'_p\times T'_q)\cap(A\times B)^*](S_p\times 1)\mid p\in P,\ q\in Q_+\}$$

Since $(A\times B)^*$ is a rational set of $A^*\times B^*$ and $S'_p\times T'_q$ is a recognizable set, their intersection is a rational set and the result follows. □

Remark 5.4. It is straightforward to see that for any $n\geqslant 2$

$$\mathrm{Rec}_n\subset\mathrm{Synch}_n\subset\mathrm{DRat}_n\subset\mathrm{Rat}_n.$$

Proposition 5.5. *It is not decidable whether a rational relation is synchronized or not.*

Proof. This proof goes as the proof that the universe problem is undecidable in Rat_2, starting with a slight reworking of the proof by Rabin and Scott that intersection is undecidable in DRat_2 (see [2]).

For any two sets of r words on A^*, $X=\{u_1,\ldots,u_r\}$ and $Y=\{v_1,\ldots,v_r\}$ let us consider the sets of pairs of words in $A^*\times B^*$, where $B=\{c,d\}$

$$U=\{(u_1,cd^{l_1}),\ \ldots,(u_r,cd^{l_r})\},$$

$$V=\{(v_1,cd^{l_1}),\ldots,(v_r,cd^{l_r})\},$$

with the exponents l_i having the following properties:
- they are all distinct,
- $l_i\geqslant 2\max(|u_i|,|v_i|)$.

As in Fischer and Rosenberg's proof [9], it is seen that U^+ and V^+ are deterministic and, thus, the relations $\overline{U^+}$ and $\overline{V^+}$ (where $\bar{Z}$ denotes the complement of Z) are rational and it is clear that the Post correspondence problem attached to X and Y has no solution if and only if $W=U^+\cap V^+$ is empty (because the l_i have been chosen all distinct), that is, if and only if $\bar{W}=\overline{U^+}\cup\overline{V^+}=A^*\times B^*$, which implies, by Proposition 5.3, that $\bar{W}$ is a synchronized rational relation.

Suppose that the correspondence problem has a solution and that $\bar{W}=\overline{U^+}\cup\overline{V^+}$ is nevertheless a synchronized rational relation. By Proposition 4.5, its complement $W=U^+\cap V^+$ is also a synchronized rational relation. If $(f,g)\in W$ then $|g|\geqslant 2|f|$ (because of the second condition on the l_i) and (f^n,g^n) belongs to W for any n: W has unbounded length difference. Since the relation W is of finite image (and its inverse is a function) there is a contradiction with Corollary 4.7.

Thus, the relation $\bar{W}$ is synchronized if and only if the Post correspondence problem attached to X and Y has no solution, which we cannot decide. □

It is known that Rec_2 is undecidable in Rat_2 (see [2]) and that DRat_2 is undecidable in Rat_2 as well [9].

Remark 5.6. We have chosen to state Theorem 5.1 under a form as simple as possible and that allows a comparison with the corresponding statement on infinite words – cf. Theorem 8.3. However, one can be much more precise in the description of the "lower" part of Rat_2 by introducing the relations with bounded length difference. Let us now briefly sketch this part of the picture, leaving the (easy) verifications to the reader.

For any nonnegative integer k let us first denote by $\mathrm{ld}_k\text{-}\mathrm{Rat}_2$ the family of rational relations with length difference smaller than or equal to k and by $\mathrm{B}_k\text{-}\mathrm{Rat}_2$ the Boolean algebra generated by $\mathrm{ld}_k\text{-}\mathrm{Rat}_2$. We have the double infinite strict hierarchy

$$\mathrm{ld}_0\text{-}\mathrm{Rat}_2 \subset \cdots \subset \mathrm{ld}_k\text{-}\mathrm{Rat}_2 \subset \mathrm{ld}_{k+1}\text{-}\mathrm{Rat}_2 \subset \cdots$$

$$\mathrm{B}_0\text{-}\mathrm{Rat}_2 \subset \cdots \subset \mathrm{B}_k\text{-}\mathrm{Rat}_2 \subset \mathrm{B}_{k+1}\text{-}\mathrm{Rat}_2 \subset \cdots$$

Now let $\mathrm{bld}\text{-}\mathrm{Rat}_2 = \bigcup_{k=0}^{\infty} \mathrm{ld}_k\text{-}\mathrm{Rat}_2$ be the family of rational relations with bounded length difference. The Boolean algebra generated by this last family is equal to $\bigcup_{k=0}^{\infty} B_k\text{-}\mathrm{Rat}_2$ and is strictly contained in Synch_2. Finally, let us denote by Fin_2 the family of relations with finite graph; then Proposition 5.7 holds.

Proposition 5.7. $\mathrm{Fin}_2 = \mathrm{Rec}_2 \cap \mathrm{bld}\text{-}\mathrm{Rat}_2$.

The characterization of the smallest Boolean algebra that contains both Rec_2 and $\mathrm{bld}\text{-}\mathrm{Rat}_2$ seems to be an open problem.

PART II: RELATIONS OF INFINITE WORDS

We now turn to infinite words and we extend some of the above properties to relations on these objects.

6. Preliminaries

The set of infinite words – often called ω-words in the literature – on an alphabet A is denoted by $A^{\mathbb{N}}$. An infinite computation c of an automaton $\mathscr{A}$ on A, $\mathscr{A} = (Q, A, E, I, T)$, is an infinite labelled path

$$c = q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \cdots$$

in the labelled graph $\mathscr{A}$. The computation c is *successful* if q_0 is an initial state and q_n belongs to T for an infinite number of indices n. This definition of successfulness is usually known as the "Büchi condition of acceptance". The *infinite behavior* of $\mathscr{A}$, denoted by $\|\mathscr{A}\|$, is the set of labels of successful computations of $\mathscr{A}$. To make the short story shorter, we take as a definition that a subset of $A^{\mathbb{N}}$ is *rational* if it is the infinite behavior of a finite automaton. The family of rational subsets of $A^{\mathbb{N}}$ is denoted by Rat $A^{\mathbb{N}}$, and by ω-Rat$_1$ if the alphabet is not specified (the reason for the index 1 will become clear in the sequel). It is known (see [6]) that, in contrast with the situation for finite words, a finite automaton is not always equivalent to a deterministic one, i.e.

$$\mathrm{DRat}\, A^{\mathbb{N}} \subset \mathrm{Rat}\, A^{\mathbb{N}} \quad \text{or} \quad \omega\text{-DRat}_1 \subset \omega\text{-Rat}_1 \tag{2}$$

and that, nevertheless, Rat $A^{\mathbb{N}}$ is an effective Boolean algebra (Büchi's Theorem [4]). On the other hand, DRat $A^{\mathbb{N}}$ is *not* a Boolean algebra.

These definitions extend, more or less directly, to relations on infinite words. Let $\mathscr{A} = (Q, A^* \times B^*, E, I, T)$ be a 2-automaton with edges labelled by elements of $A^* \times B^*$. Let

$$c = q_0 \xrightarrow{(x_1, y_1)} q_1 \xrightarrow{(x_2, y_2)} q_2 \cdots$$

be an infinite computation of $\mathscr{A}$; we keep the same Büchi condition of acceptance: c is *successful* if q_0 is an initial state and q_n belongs to T for an infinite number of indices n. The *label* of c is the pair $(x_1 x_2 \ldots, y_1 y_2 \ldots)$. The problem is that such a label is not necessarily an element of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ for it may happen that the x_i's (or the y_i's) are all equal to 1 from a certain rank on. We shall say that c is *admissible* if $x_1 x_2 \ldots$ and $y_1 y_2 \ldots$ are both infinite words. The following proposition shows that one can indeed get (effectively) rid of this restriction.

Proposition 6.1. *Given a 2-automaton $\mathscr{A}$, it is possible to effectively construct a 2-automaton $\mathscr{B}$ such that every infinite successful computation in $\mathscr{B}$ is admissible and such that the set of labels of successful and admissible infinite computations of $\mathscr{A}$ and $\mathscr{B}$ are equal.*

There are several possible proofs for this proposition (see also [13]). The first one we give just below makes use of Büchi's Theorem. It is not valid for deterministic 2-automata for which we shall give another proof with a more explicit construction (see Appendix A).

Proof. Let $\mathscr{A} = (Q, A^* \times B^*, E, I, T)$ be a finite 2-automaton. Let X be a set in one-to-one correspondence with E: for every x in X, $\varphi(x)$ is the label of the edge x. One derives immediately from $\mathscr{A}$ the 1-automaton $\mathscr{A}' = (Q, X, E', I, T)$ with the same underlying graph $[(p, x, q) \in E' \Leftrightarrow (p, \varphi(x), q) \in E]$. There is a 1–1 correspondence between the computations (finite or infinite) of $\mathscr{A}$ and $\mathscr{A}'$, and if c and c' are corresponding computations then $|c| = \varphi(|c'|)$. Let L' be the infinite behavior of $\mathscr{A}'$: $L' \in \mathrm{Rat}\, X^{\mathbb{N}}$. Let $Y = \{x \in X \mid \varphi(x) \in A^+ \times 1\}$ and let $Z = \{x \in X \mid \varphi(x) \in 1 \times B^+\}$. Then $\varphi(L' \setminus (X^* Y^\omega \cup X^* Z^\omega))$ is exactly the set of labels of admissible and successful infinite computations of $\mathscr{A}$. Then $L' \setminus (X^* Y^\omega \cup X^* Z^\omega)$ is a rational set of $X^{\mathbb{N}}$ and the infinite behavior of a 1-automaton $\mathscr{B}' = (P, X, F', J, S)$ effectively constructible from $\mathscr{A}'$ (by

Büchi's Theorem). The "image" of $\mathscr{B}'$ by φ yields the 2-automaton $\mathscr{B}=(P, A^* \times B^*, F, J, S)$ that answers the problem. $\square$

From now on we assume that in any 2-automaton every successful infinite computation is admissible. The *infinite behavior* $\|\mathscr{A}\|$ of a 2-automaton $\mathscr{A}$ is the set of labels of successful infinite computations of $\mathscr{A}$.

A relation of infinite words – also called ω-relation – is a subset of $A^{\mathbb{N}} \times B^{\mathbb{N}}$. We take as a definition that such an ω-relation is *rational* if it is the infinite behavior of a finite 2-automaton. The equivalence of this definition with a definition by means of ω-rational operations, as well as a number of extensions of properties of rational relations to rational ω-rational relations can be found in [13]. The set of rational relations of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ is denoted by $\mathrm{Rat}(A^{\mathbb{N}} \times B^{\mathbb{N}})$, by ω-Rat_2 if the alphabets are not specified. In the sequel, all the (2-) automata considered will be finite and thus simply called (2-)automata.

Since every element $(a_1 a_2 \ldots, b_1 b_2 \ldots)$ of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ – with a_i in A and b_j in B – may be written as the infinite product $((a_1, b_1)(a_2, b_2)\cdots)$, it holds that

$$A^{\mathbb{N}} \times B^{\mathbb{N}} = (A \times B)^{\mathbb{N}},$$

the counterpart of which in the domain of finite words is

$$A^* \times B^* = (A \times B)^* [(A^* \times 1) \cup (1 \times B^*)].$$

On the other hand, the inclusion

$$\mathrm{Rat}(A \times B)^{\mathbb{N}} \subset \mathrm{Rat}(A^{\mathbb{N}} \times B^{\mathbb{N}})$$

is strict. Somehow this inclusion is all what this part of the paper is about.

Definition 6.2. A rational relation of infinite words is said to be *synchronized* if it is the infinite behavior of a letter-to-letter 2-automaton.

The set of synchronized rational relations, equal by definition to $\mathrm{Rat}(A \times B)^{\mathbb{N}}$, is also denoted by $\mathrm{Synch}(A^{\mathbb{N}} \times B^{\mathbb{N}})$ in order to make the whole set of notations more consistent, and by ω-Synch_2 if the alphabets are not specified. It seems that synchronized rational relations have been considered first by Büchi as the realization of his "sequential calculus". They have been denoted by SEQ (or ω-SEQ) in [23] and by ω-BÜC in [24].

7. Rational ω-relations with bounded delay

If the length difference of a relation of finite words has no equivalent for a relation of infinite words, the *delay* is defined in the same way for infinite as for finite computations: if c is an infinite computation of $\mathscr{A}$

$$\mathrm{del}(c) = \sup\{\mathrm{del}(c, n) \mid n \geqslant 1\}$$

and the ω-delay of the automaton $\mathscr{A}$ is the supremum of the delay of successful infinite computations of $\mathscr{A}$.

Definition 7.1. A rational relation of infinite words is said to have a *bounded delay* if it is the infinite behavior of a 2-automaton with bounded delay.

The notion of trimmed automaton has to be redefined in connection with infinite behavior: a 2-automaton is *ω-trimmed* if (a) every state is accessible, (b) every state is ω-coaccessible, i.e. is the starting state of an infinite computation going infinitely often through the set of terminal states. As in the case of 1-automata (see [19]) we have Proposition 7.2.

Proposition 7.2. *Given a 2-automaton $\mathscr{A}$, it is possible to construct effectively an ω-trimmed 2-automaton $\mathscr{B}$ that has the same infinite behavior as $\mathscr{A}$.*

Proof. Let $\mathscr{A}=(Q, A^* \times B^*, E, I, T)$ be a 2-automaton and let P be the set of accessible and ω-coaccessible states of $\mathscr{A}$. Let $\mathscr{B}=(P, A^* \times B^*, F, I \cap P, T \cap P))$, where $F = E \cap (P \times A^* \times B^* \times P)$. Then, clearly, it holds that $\|\mathscr{B}\| \subseteq \|\mathscr{A}\|$.

Conversely, let $(x_1 x_2 \ldots, y_1 y_2 \ldots) \in \|\mathscr{A}\|$. Thus, there exists a path $c = p_0 \rightarrow p_1 \rightarrow p_2 \cdots$ with $p_0 \in I$, going infinitely often through T and labelled by $(x_1 x_2 \ldots, y_1 y_2 \ldots)$. Every p_i in c is accessible and is the beginning of an infinite path that goes infinitely often through T and, thus, belongs to P: c is a computation of $\mathscr{B}$, a successful one. $\square$

It is clear that an ω-trimmed automaton is trimmed and that there is no difference between the delay and the ω-delay for ω-trimmed automata. Another formulation of Lemma 3.1 is thus: "*An ω-trimmed 2-automaton has a bounded ω-delay if and only if the length difference of the label of every loop is* 0".

Remark 7.3. We have defined 2-automata with bounded delay. We could have distinguished two different kinds of 2-automata that do not have bounded delay. One would say that the ω-delay of a 2-automaton is *unbounded* if the delay of every successful infinite computation is *finite*, but there is *no bound* on the set of the delays of these computations. On the other hand, one would say that the ω-delay of a 2-automaton is *infinite* if there exists *one successful infinite* computation which has an *infinite* delay. One can decide between an unbounded and an infinite ω-delay for a 2-automaton. The ω-delay is infinite if and only if there exists a loop, the label of which has length difference different from 0, which contains a terminal state.

As a corollary of Lemma 3.1 and a consequence of the classical construction of the ω-rational expression of the infinite behavior of a finite automaton (see [7]), we have Corollary 7.4.

Corollary 7.4. *A rational relation of infinite words with bounded delay may be written as a finite union $R = \bigcup_i S_i T_i^{\omega}$, where S_i is a rational relation of finite words with bounded delay and T_i is a length-preserving rational relation of finite words.*

Example 7.5. Let R be the relation that maps every word of the form a^*b^ω to the word b^ω. The graph of R is $(a, b)^*(b, b)^\omega$ and is realized by a 0-delay 2-automaton.

Example 7.6. There are rational relations of infinite words which do not have a bounded delay. Let R be the relation which maps every word of the form $a^{2n_1}ba^{2n_2}b\ldots$ to the word $a^{n_1}ba^{n_2}b\ldots$. The graph of R is $[(a^2, a)^*(b, b)]^\omega$. It is not difficult to check that R cannot be realized by a 2-automaton the loops of which are length-preserving. (Proposition 7.8 will give other evidences of rational relations with unbounded delay.)

Remark 7.7. In opposition to the case of finite words (Proposition 3.2), it is not true that every 2-automaton that realizes a rational relation with bounded delay has a bounded delay. For instance, the relation R of Example 7.5 is also given by the expression $(a, 1)^*(b, b)^\omega$ which corresponds to a 2-automaton with unbounded delay.

Proposition 7.8 (Frougny [10]). *A rational relation R of infinite words has a bounded delay if and only if R is a synchronized rational relation.*

The overall principle of the proof is the same as in the finite case: in view of Corollary 7.4, the idea is to postpone the length difference by replacing a product $(u, 1)L$, where L is a length-preserving rational relation with a finite union of $L_i(u_i, 1)$. The problem of the "infinite" postponement is solved by the fact that there are only a finite number of words of length k, as expressed by the following lemma and its proof.

Lemma 7.9 (Resynchronization lemma for ω-words). *Let T be a rational set of $(A \times B)^*$ and $(u, 1)$ be an element of $A^* \times 1$. Then there exists a finite number of T_i and S_i of* $\mathrm{Rat}(A \times B)^*$ *such that* $(u, 1)T^\omega = \bigcup_i S_i T_i^\omega$.

Proof. Let $k = |u|$; from Lemma 2.3 follows that for every $v \in A^k$ there exist a finite number of rational subsets V_j of $(A \times B)^*$ and a finite number of words v_j of A^k such that

$$(v, 1)T = \bigcup_j V_j(v_j, 1). \tag{3}$$

Let $\mathscr{S}$ be the generalized 2-automaton $\mathscr{S} = (A^k, \mathrm{Rat}(A \times B)^*, E, u, A^k)$ the edges of which are defined by (3): for every "state" v there are edges from v to v_j labelled by V_j.

We call $\mathscr{S}$ *generalized* because the edges are labelled by *subsets* of the input monoid instead by elements of the input monoid. Since rational sets are closed by substitutions, the behavior (finite or infinite) of a generalized finite automaton with rational labels is a rational set. The behavior of $\mathscr{S}$ is, thus, of the prescribed form and $\|\mathscr{S}\| = \bigcup_i S_i T_i^\omega$ with S_i and T_i in $\mathrm{Rat}(A \times B)^*$.

It just remains to show that the infinite behavior of $\mathscr{D}$ is equal to $(u, 1)T^\omega$. Let $s\in(u, 1)T^\omega$. Then s may be written as $s=(u, 1)t_1t_2\ldots$. By definition of $\mathscr{D}$, there exists an infinite successful computation

$$c=f_0 \xrightarrow{U_1} f_1 \xrightarrow{U_2} f_2\cdots$$

with the following properties: $f_0=u$, and $\forall i\in\mathbf{N}$, $(f_{i-1}, 1)t_i=u_i(f_i, 1)\in U_i(f_i, 1)$. Then, clearly, $s=u_1u_2\ldots\in U_1U_2\ldots$.

Conversely, let $w=v_1v_2\ldots$ be a word that belongs to the subset of $A^{\mathbf{N}}\times B^{\mathbf{N}}$ defined by an infinite successful computation of $\mathscr{D}$

$$d=u \xrightarrow{V_1} g_1 \xrightarrow{V_2} g_2\cdots$$

Then using (3) again, the second component of w may be "shifted" by k positions, giving rise to a factorization $w=(u, 1)t_1t_2\ldots\in(u, 1)T^\omega$. $\square$

Proof of Proposition 7.8. Clearly, a rational subset of $(A\times B)^{\mathbf{N}}$ has a bounded delay (a delay 0 indeed). Conversely, let R be a rational relation with bounded delay. From Corollary 7.4, $R=\bigcup_i S_iT_i^\omega$, where the S_i's are rational relations of $A^*\times B^*$ with bounded delay, and the T_i's belong to $\mathrm{Rat}(A\times B)^*$. By Proposition 2.1, every S_i is a finite union $S_i=\bigcup_j H_{i,j}u_{i,j}$, where $H_{i,j}\in\mathrm{Rat}(A\times B)^*$ and $|u_{i,j}|\leqslant k$, for a certain integer k. Then $R=\bigcup_i(\bigcup_j H_{i,j}u_{i,j}T_i^\omega)$ and the conclusion follows from the resynchronization lemma. $\square$

This result may be rephrased as Corollary 7.10.

Corollary 7.10. *A rational relation of infinite words with bounded delay may be written as a finite union* $R=\bigcup_i S_iT_i^\omega$, *where* S_i *and* T_i *are length-preserving rational relations of finite words.*

The property stated in Proposition 7.8 has been independently proved – for a restricted class of rational relations with bounded delay – in [17]. From Proposition 7.8 and the definition of synchronized rational relations then follows Corollary 7.11.

Corollary 7.11. *The rational relations of infinite words with bounded delay form an (effective) Boolean algebra.*

8. On the geography of ω-Rat$_2$

The sequence of inclusions in Theorem 5.1 that holds for relations of finite words become somewhat more intricate for relations of infinite words because of the distinction that has to be made between the relations defined by deterministic and nondeterministic automata, even in the case of 1-automata [cf. (2)]. The deterministic rational relations of infinite words, denoted by $\mathrm{DRat}(A^{\mathbb{N}} \times B^{\mathbb{N}})$ or ω-DRat_2, are those realized by deterministic 2-automata – with the same definition as in Section 5 for deterministic rational relations of finite words (and even simpler since there is no need for any endmarker!). To make this definition sound, we have also to show that we can choose the (deterministic) automaton in such a way that all successful computations are admissible.

Proposition 8.1. *Given a deterministic 2-automaton $\mathscr{A}$, it is possible to effectively construct a deterministic 2-automaton $\mathscr{B}$ such that*

(i) *every infinite successful computation in $\mathscr{B}$ is admissible,*

(ii) *the set of labels of successful and admissible infinite computations of $\mathscr{A}$ and $\mathscr{B}$ are equal.*

The proof of this proposition is postponed to Proposition A.3.

We denote by $\mathrm{DSynch}(A^{\mathbb{N}} \times B^{\mathbb{N}})$, or by ω-DSynch_2, the synchronized rational relations realized by deterministic letter-to-letter 2-automata (deterministic as 1-automata over the alphabet $A \times B$). Finally, let us define (see also [24]) the equivalent for infinite words of recognizable relations (of finite words).

Definition 8.2. A *componentwise rational relation* of infinite words is a finite union of cartesian products of the form $S \times T$, with S in $\mathrm{Rat}\, A^{\mathbb{N}}$ and T in $\mathrm{Rat}\, B^{\mathbb{N}}$.

The corresponding family of relations is denoted by $\mathrm{CwRat}(A^{\mathbb{N}} \times B^{\mathbb{N}})$, or ω-CwRat_2. It is easy to see that ω-CwRat_2 is an effective Boolean algebra contained in ω-Rat_2.

With these definitions, and notations, the following theorem then holds.

Theorem 8.3.

$$\begin{array}{ccccc} & & \omega\text{-DSynch}_2 & \subset & \omega\text{-DRat}_2 \\ & & \cap & & \cap \\ \omega\text{-CwRat}_2 & \subset & \omega\text{-Synch}_2 & \subset & \omega\text{-Rat}_2 \end{array} \tag{4}$$

The vertical inclusions in (4) are obvious; the fact they are strict follows from (2). The identity relation belongs to ω-$\mathrm{DSynch}_2 \backslash \omega$-$\mathrm{CwRat}_2$; the relation $R = [(aa, a) \cup (b, b)]^{\omega}$ belongs to ω-$\mathrm{DRat}_2 \backslash \omega$-$\mathrm{Synch}_2$ (this is a small exercise left to the reader), showing that the horizontal inclusions – that remain to be proved – are strict. The inclusion ω-$\mathrm{DSynch}_2 \subset \omega$-$\mathrm{DRat}_2$ is established as was $\mathrm{Synch}_2 \subset \mathrm{DRat}_2$, and is simpler since there is no endmarker. The last one is stated in the following.

Proposition 8.4. *A componentwise rational relation of infinite words is a synchronized relation of infinite words.*

This statement has been independently announced in [24].

Proof of Proposition 8.4. With no loss of generality, we assume that $R = S \times T$ with $S \in \mathrm{Rat}\, A^{\mathbb{N}}$ and $T \in \mathrm{Rat}\, B^{\mathbb{N}}$. We have

$$S \times T = (S \times B^{\mathbb{N}}) \cap (A^{\mathbb{N}} \times T).$$

Since ω-Synch$_2$ is closed under intersection, it is enough to consider relations of the form $S \times B^{\mathbb{N}}$. Let $\mathscr{A} = (Q, A, E, I, T)$ be an automaton which recognizes S. Then we construct a letter-to-letter 2-automaton $\mathscr{A}'$ which recognizes $S \times B^{\mathbb{N}}$ in the following way: $\mathscr{A}' = (Q, A \times B, E', I, T)$, where

$$E' = \{(p, (a, b), q) \mid (p, a, q) \in E, \forall b \in B\}. \qquad \square$$

The proof of Theorem 8.3 is, thus, complete.

Proposition 8.5. *Let S be in* $\mathrm{DRat}\, A^{\mathbb{N}}$ *and T in* $\mathrm{DRat}\, B^{\mathbb{N}}$. *Then $S \times T$ belongs to* $\mathrm{DSynch}(A^{\mathbb{N}} \times B^{\mathbb{N}})$.

Proof. The construction given in the proof of the Proposition 8.4 gives a deterministic letter-to-letter 2-automaton (that is a deterministic 1-automaton over $A \times B$) when applied to a deterministic automaton that recognizes S (or T). Now, since $\mathrm{DSynch}(A^{\mathbb{N}} \times B^{\mathbb{N}}) = \mathrm{DRat}(A \times B)^{\mathbb{N}}$, ω-DSynch$_2$ is closed under intersection (see [19]) and, thus, the result follows. $\square$

Remark 8.6. Let R be a nonempty relation of finite words equal to $S \times T$ with $S \in \mathrm{Rat}\, A^*$ and $T \in \mathrm{Rat}\, B^*$. Then $R^{\omega} = S^{\omega} \times T^{\omega}$ belongs to $\mathrm{CwRat}(A^{\mathbb{N}} \times B^{\mathbb{N}})$ and, thus, to $\mathrm{Rat}(A \times B)^{\mathbb{N}}$. On the other hand, it is not true that if R is a recognizable relation of finite words then R^{ω} is in $\mathrm{Rat}(A \times B)^{\mathbb{N}}$. [Consider, for instance, the relation $R = (aa, a) \cup (b, b)$.]

Theorem 8.7. ω-DSynch$_2 = \omega$-Synch$_2 \cap \omega$-DRat$_2$.

The inclusion from left to right is obvious. The reverse inclusion will be shown by means of a topological argument.

Recall that the set $A^{\mathbb{N}}$ is equipped with the product topology, induced by the discrete topology on A. In that topology the open sets are the sets of the form $XA^{\mathbb{N}}$ with $X \subset A^+$ (see [25] or [19] for instance). Thus, the open sets of $(A \times B)^{\mathbb{N}}$ are of the form $Z(A \times B)^{\mathbb{N}}$, with $Z \subset (A \times B)^+$.

The open sets of the product $A^{\mathbb{N}} \times B^{\mathbb{N}}$ are of the form $U \times V$, where U is an open set of $A^{\mathbb{N}}$ and V is an open set of $B^{\mathbb{N}}$; therefore, are of the form $XA^{\mathbb{N}} \times YB^{\mathbb{N}}$, with $X \subset A^+$ and $Y \subset B^+$. Since

$$\begin{aligned}&[XA^{\mathbb{N}} \times B^{\mathbb{N}}] \cap [A^{\mathbb{N}} \times YB^{\mathbb{N}}]\\ &\quad = [(X \times 1)(A^* \times B^*) \cap (1 \times Y)(A^* \times B^*) \cap (A \times B)^*](A \times B)^{\mathbb{N}}\end{aligned}$$

the topologies of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ and of $(A \times B)^{\mathbb{N}}$ coincide.

The basic relationship between topology and rational sets of infinite words is given by the following proposition.

Proposition 8.8 (Landweber [16]). *A rational subset X of $A^{\mathbb{N}}$ is recognized by a deterministic automaton if and only if X is a countable intersection of open sets.*

For rational ω-relations the above condition is necessary as expressed by the following proposition.

Proposition 8.9. *The infinite behavior of a deterministic 2-automaton is a countable intersection of open sets.*

The proof consists in two lemmas.

An element $w \in A^* \times B^*$ is said to be a *prefix* of $s \in A^{\mathbb{N}} \times B^{\mathbb{N}}$ if $w=(f, g)$, $s=(u, v)$ and f is a prefix of u, g is a prefix of v. Let $L \in A^+ \times B^+$. The set of elements of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ having infinitely many prefixes in L is denoted by $\vec{L}$.

As in the case of deterministic 1-automata, and because in a deterministic 2-automaton any pair of infinite words determines a unique computation that can be viewed as a unique infinite word, the following holds.

Lemma 8.10. *Let $\mathscr{A}$ be a deterministic 2-automaton. Then $\|\mathscr{A}\| = \overrightarrow{|\mathscr{A}|}$.*

Proof. Let $\mathscr{A}=(Q_A \cup Q_B, (A \times 1) \cup (1 \times B), E, \{i\}, T)$ be a deterministic 2-automaton. If $s \in \|\mathscr{A}\|$, then it is the label of an infinite path $(q_0, s_1, q_1)(q_1, s_2, q_2)\ldots$, with $q_0=i$ and $s_k \in (A \times 1) \cup (1 \times B)$, and such that there exsists a subsequence $n_0 < n_1 < n_2 < \cdots$ such that $q_{n_0}, q_{n_1}, \ldots \in T$. Hence, the words $s_1 \ldots s_{n_k} \in |\mathscr{A}|$ and are prefixes of s.

Conversely, let $s \in \overrightarrow{|\mathscr{A}|}$. Then s has infinitely many prefixes in $|\mathscr{A}|$. Since $\mathscr{A}$ is deterministic, s is the label of an infinite path starting in $\{i\}$ and going infinitely often through T. Thus, $s \in \|\mathscr{A}\|$. $\square$

As in the case of ω-words on one tape we have Lemma 8.11.

Lemma 8.11. *A subset R of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ is a countable intersection of open sets if and only if there exists a subset H of $A^* \times B^*$ such that $R=\vec{H}$.*

Proof. (1) Let R be a countable intersection of open sets of $A^{\mathbb{N}} \times B^{\mathbb{N}}$. Since the topologies of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ and of $(A \times B)^{\mathbb{N}}$ are identical, R can be viewed as a countable intersection of open sets of $(A \times B)^{\mathbb{N}}$; thus, there exists (by [16]) a set H in $(A \times B)^*$ such that $R=\vec{H}$.

(2) Conversely, let $R=\vec{H}$ and let, for every $n \geqslant 0$, $H_n=\{(u, v) \in H \mid |u|+|v|>n\}$. We show that

$$\vec{H} = \bigcap_{n \geqslant 0} [H_n(A^{\mathbb{N}} \times B^{\mathbb{N}})].$$

Let $s \in \vec{H}$. Then for every $n \geqslant 0$, s has a prefix (u, v) in H with $|u|+|v|>n$. Thus, $s \in \bigcap_{n \geqslant 0}[H_n(A^{\mathbb{N}} \times B^{\mathbb{N}})]$. Conversely, if $s \in \bigcap_{n \geqslant 0}[H_n(A^{\mathbb{N}} \times B^{\mathbb{N}})]$, then for every n, s is of the form $s=(u, v)(x, y)$, with $(u, v) \in H_n$. Hence, s has infinitely many prefixes in H and $s \in \vec{H}$. Now, for every $Z \subset A^* \times B^*$,

$$Z(A^{\mathbb{N}} \times B^{\mathbb{N}}) = \bigcup_{(u,v) \in Z} (u, v)(A^{\mathbb{N}} \times B^{\mathbb{N}})$$

and is open as union of (elementary) open sets. □

Proof of Proposition 8.9. Let R be the infinite behavior of a deterministic 2-automaton. By Lemma 8.10, R is equal to $\vec{H}$, where H is the finite behavior of the automaton. By Lemma 8.11, the result follows. □

Remark 8.12. It is not true that, if a rational subset of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ is a countable intersection of open sets, then it is deterministic. Let $K = \{(a^{2n}b, a^n b) \mid n \in \mathbb{N}\} \cup \{(a^{2n+1}b, a^{2n+1}b) \mid n \in \mathbb{N}\}$ and $H = K^*$. Then H is a nondeterministic rational relation. Now take $R = K^{\omega}$. It is easy to check that R is a nondeterministic rational ω-relation and that R is equal to $\vec{H}$. By Lemma 8.11, R is a countable intersection of open sets.

We now come to our point.

Proof of Theorem 8.7. Let R be a deterministic rational relation of $A^{\mathbb{N}} \times B^{\mathbb{N}}$. Then R is a countable intersection of open sets of $A^{\mathbb{N}} \times B^{\mathbb{N}}$ and, thus, of open sets of $(A \times B)^{\mathbb{N}}$, since the topologies are identical. If in addition R is synchronized, R is a subset of $\mathrm{Rat}(A \times B)^{\mathbb{N}}$. Now, by Proposition 8.8, R is a deterministic rational relation of $(A \times B)^{\mathbb{N}}$ and, thus, a deterministic synchronized rational relation. □

In order to complete the picture given by Theorem 8.3, we have to consider the problem of decidability of subfamilies within ω-Rat_2. The undecidability results for relations of finite words will carry over to relations of infinite words by means of the notion of *ω-completion* of a relation of finite words.

Definition 8.13. If R is a relation of finite words, the *ω-completion* of R is the relation of infinite words $R(\$^{\omega}, \$^{\omega})$, where $\$$ is a new letter.

Clearly, the ω-completion of a rational relation is a rational ω-relation. Conversely, if T is a rational ω-relation which is an ω-completion, T can be written as a finite union $T = \bigcup X_i Y_i^{\omega}$, where $Y_i \in \mathrm{Rat}(\$^+ \times \$^+)$ and $X_i \subset (A^* \times B^*)(\$^* \times \$^*)$. Let π be the morphism which deletes $\$$. Then T is the ω-completion of the rational relation $R = \bigcup \pi(X_i)$. This property extends to deterministic relations.

Proposition 8.14. *A rational relation of finite words is deterministic if and only if its ω-completion is a deterministic rational ω-relation.*

We give the corresponding construction in Appendix A. Together with the result of [9], already quoted in Section 5, that DRat_2 is undecidable in Rat_2, Proposition 8.14 yields the following proposition.

Proposition 8.15. *It is not decidable whether a rational ω-relation is deterministic or not.*

Coming back to synchronized rational relations, we have the following result.

Lemma 8.16. *A relation of finite words is a left-synchronized rational relation if and only if its ω-completion is a synchronized rational relation of infinite words.*

Proof. If R is left-synchronized then it is equal to a finite union,

$$R=[\cup_i L_i(S_i\times 1)]\bigcup[\cup_j M_j(1\times T_j)],$$

with L_i and M_j in $\mathrm{Rat}(A\times B)^*$ and S_i in $\mathrm{Rat}\,A^*$, T_j in $\mathrm{Rat}\,B^*$. Let $H_i=\{(f,\$^{|f|})\mid f\in S_i\}$; then H_i belongs to $\mathrm{Rat}(A\times\$)^*$. Similarly, $K_j=\{(\$^{|f|},f)\mid f\in T_j\}$ is an element of $\mathrm{Rat}(\$\times B)^*$. Thus,

$$R(\$^\omega,\$^\omega)=[\cup_i L_iH_i(\$,\$)^\omega]\bigcup[\cup_j M_jK_j(\$,\$)^\omega],$$

which is a synchronized rational ω-relation.

Conversely, let $R\subset A^*\times B^*$ be a relation such that its ω-completion $R(\$^\omega,\$^\omega)$ is a synchronized rational relation. Then $R(\$^\omega,\$^\omega)$ can be written as a finite union,

$$R(\$^\omega,\$^\omega)=[\cup_i L_iH_i(\$,\$)^\omega]\bigcup[\cup_j M_jK_j(\$,\$)^\omega],$$

where L_i and M_j are in $\mathrm{Rat}(A\times B)^*$, H_i is in $\mathrm{Rat}(A\times\$)^*$ and K_j is in $\mathrm{Rat}(\$\times B)^*$. Thus, when the $\$$'s are erased in the H_i's and K_j's, and the factor $(\$^\omega,\$^\omega)$ is deleted, R is clearly a left-synchronized rational relation. $\square$

Together with Proposition 5.5, this result yields the following proposition.

Proposition 8.17. *It is not decidable whether a rational relation of infinite words is synchronized or not.*

On the other hand, and since it is decidable whether an element of $\mathrm{Rat}\,A^{\mathbb{N}}$ is deterministic or not [16], Proposition 8.18 follows.

Proposition 8.18. *It is decidable whether a synchronized rational relation of infinite words is deterministic or not.*

Appendix A: constructions on automata

A.1. Covering of automata

The aim of this section is just to adapt or extend to automata the notion of *covering* as defined by Stallings [22] for graphs.

A.1.1. Morphism of automata

Stallings's definition of graphs can first be slightly transformed as follows. A *graph* $\mathscr{G}$ is a pair (Q, E) of two sets: Q is the set of *vertices*, E the set of *edges*, equipped with two mappings $\iota: E \to Q$ and $\tau: E \to Q$. The vertices $\iota(e)$ and $\tau(e)$ are, respectively, the *origin* and the *end* of the edge e.

A *morphism* from a graph $\mathscr{H} = (R, F)$ into a graph $\mathscr{G} = (Q, E)$ is a pair of mappings (both denoted by φ) $\varphi: R \to Q$ and $\varphi: F \to E$, with the property that

$$\iota \circ \varphi = \varphi \circ \iota, \tag{5}$$

$$\tau \circ \varphi = \varphi \circ \tau. \tag{6}$$

These two conditions imply that the image of a path in $\mathscr{H}$ is a path in $\mathscr{G}$.

A *labelled* graph $\mathscr{G} = (Q, M, E)$ is a graph (Q, E), together with a third mapping ε from E into a monoid M: $\varepsilon(e)$ is the *label* of the edge e.

A morphism φ from $\mathscr{H} = (R, M, F)$ into $\mathscr{G} = (Q, M, E)$ is a pair of mappings as above that satisfy also, in addition to (5) and (6),

$$\varepsilon \circ \varphi = \varepsilon. \tag{7}$$

This implies that the label of a path is the same as the label of the image of that path.

A morphism φ from an automaton $\mathscr{B} = (R, M, F, J, U)$ into an automaton $\mathscr{A} = (Q, M, E, I, T)$ satisfy, in addition to (5)–(7), the two conditions

$$\varphi(J) \subseteq I, \tag{8}$$

$$\varphi(U) \subseteq T. \tag{9}$$

That is, the image of a successful (infinite successful) path in $\mathscr{B}$ is a successful (infinite successful) path in $\mathscr{A}$ – and with the same label. In particular, $|\mathscr{B}| \subseteq |\mathscr{A}|$ and $\|\mathscr{B}\| \subseteq \|\mathscr{A}\|$.

A.1.2. Covering of automata

For every vertex q of a graph $\mathscr{G} = (Q, E)$ let us denote by $\mathrm{Out}_{\mathscr{G}}(q)$ the set[1] of edges of $\mathscr{G}$ the origin of which is q; that is, edges that are "going out" of q:

$$\mathrm{Out}_{\mathscr{G}}(q) = \{e \in E \mid \iota(e) = q\}.$$

[1] denoted by $\mathrm{St}(q, \mathscr{G})$ and called "star" of q in $\mathscr{G}$ in [22].

For a morphism φ from $\mathscr{H}=(R, F)$ into $\mathscr{G}=(Q, E)$ and for every vertex r in $\mathscr{H}$, let us denote by φ_r the restriction of $\varphi: F \to E$ to $\mathrm{Out}_{\mathscr{H}}(r)$

$$\varphi_r : \mathrm{Out}_{\mathscr{H}}(r) \to \mathrm{Out}_{\mathscr{G}}(\varphi(r)).$$

By definition, the morphism φ is a *covering* (resp. an *immersion*) if, for every r in R, φ_r is bijective (resp. injective).

The property of coverings that will be used here is the possibility of *lifting the paths*: if $\varphi: \mathscr{H} \to \mathscr{G}$ is a covering, r a vertex in $\mathscr{H}$, and c a path in $\mathscr{G}$ such that $\iota(r)=\iota(c)$, then there exists a *unique* path d in $\mathscr{H}$ with origin r and such that $\varphi(d)=c$.

A morphism φ from an automaton $\mathscr{B}=(R, M, F, J, U)$ into an automaton $\mathscr{A}=(Q, M, E, I, T)$ is a *covering* if

(i) φ is a covering of (R, F) onto (Q, E),

(ii) $\forall i \in I$ there exists a unique $j \in J$ such that $\varphi(j)=i$,

(iii) $\forall t$, $\varphi^{-1}(t) \subset U$, i.e. $\varphi^{-1}(T)=U$.

The immediate consequence of that definition is the following proposition.

Proposition A.1. *If $\varphi: \mathscr{B} \to \mathscr{A}$ is a covering then for every successful path c in $\mathscr{A}$ there exists a unique successful path d in $\mathscr{B}$ such that $\varphi(d)=c$. In particular, $|\mathscr{B}|=|\mathscr{A}|$ and $\|\mathscr{B}\|=\|\mathscr{A}\|$.*

A morphism $\varphi: \mathscr{B} \to \mathscr{A}$ is an *immersion* if

(i) $\varphi: (Q, E) \to (R, F)$ is an immersion,

(ii) $\forall i \in I$ $\mathrm{card}(\varphi^{-1}(i) \cap I) \leqslant 1$.

If φ is an immersion, in addition to $|\mathscr{B}| \subseteq |\mathscr{A}|$, it holds that the multiplicity in $\mathscr{B}$ of every element in $|\mathscr{B}|$ is smaller than or equal to its multiplicity in $|\mathscr{A}|$.

We shall say that $\mathscr{B}$ is a *subautomaton* of $\mathscr{A}$ if $\mathscr{B}$ has the same set of states as $\mathscr{A}$ and if the identity mapping on this set is (a morphism and) an immersion, that is if $\mathscr{B}$ is obtained from $\mathscr{A}$ by deleting edges and/or suppressing the quality of being initial or terminal to certain states. Finally, it will be convenient to say that $\mathscr{B}$ *covers* $\mathscr{A}$ or is a *covering* of $\mathscr{A}$ (resp. is an *immersion* in $\mathscr{A}$) if there exists a morphism $\varphi: \mathscr{B} \to \mathscr{A}$ that is a covering (resp. an immersion).

A.2. The "resynchronization" of an automaton with bounded delay

We sketch here an algorithm that produces a letter-to-letter 2-automaton (with terminal function) equivalent to a given automaton with bounded delay.

We, thus, start with any automaton $\mathscr{A}$ on $A^* \times B^*$: $\mathscr{A}=(Q, A^* \times B^*, E, I, T)$. The first step of the algorithm consists in deciding whether $\mathscr{A}$ has a bounded delay or not (as proved in Lemma 3.1, this decision is achieved by computing the length difference of the label of every simple loop of $\mathscr{A}$).

Let us now call *balance* of (f, g) the signed difference $|f|-|g|$ (i.e. what we have called length difference above is the absolute value of the balance). The second step consists in transforming $\mathscr{A}$ into a 2-automaton $\mathscr{B}$ with the property that every state can be given a rank (an integer, either positive or negative) which is the *unique* balance of the label of *any* path from *any* initial state to that state.

The canonical (or minimal) automaton that achieves this property is obtained from a covering of $\mathscr{A}$ where the preimage of any state p is $p \times \mathbf{Z}$ and where the preimage of any edge $e=(p, u, q)$ is the set of edges $((p, z), u, (q, z+d))$ for every z in $\mathbf{Z}$ and where d is the balance of the label $u \in A^* \times B^*$ of e. More formally, let $\mathscr{A}'$ be the (infinite) automaton

$$\mathscr{A}'=(Q \times \mathbf{Z}, A^* \times B^*, E', I \times \{0\}, T \times \mathbf{Z}),$$

with

$$E'=\{((p, z), (f, g), (q, z+d)) \mid \forall (p, (f, g), q) \in E \ \forall z \in \mathbf{Z} \ d=|f|-|g|\}.$$

The *rank* of the state (p, z) is precisely z; every initial state of $\mathscr{A}'$ has rank 0; if t is a terminal state of $\mathscr{A}$, (t, z) is a terminal state of $\mathscr{A}'$ for any z in $\mathbf{Z}$.

Let $\mathscr{B}$ be the accessible part of $\mathscr{A}'$; it should be clear that $\mathscr{B}$ is finite since $\mathscr{A}$ has a bounded delay, and effectively constructible.

From now on, we, thus, assume that $\mathscr{A}$ is "ranked". $\mathscr{A}$ is a letter-to-letter 2-automaton – or, more accurately, an automaton on $(A \times B)^*$ – if and only if the rank of every state is 0.

Let us suppose that not every state of $\mathscr{A}$ has rank 0, i.e. there are some states with strictly positive rank, and let r be the supremum of the rank of states. The third step of the algorithm consists in building a covering $\mathscr{B}$ of $\mathscr{A}$, which is a ranked automaton with a supremum of rank equal to $r-1$. Iterating this procedure r times, we, thus, get an automaton having no state of strictly positive rank. A symmetric procedure (the symmetry exchanges the two tapes) will then give an automaton having no state of strictly negative rank either; hence, an automaton on $(A \times B)^*$.

Let $\mathscr{A}=(Q, A^* \times B^*, E, I, \omega)$ be a ranked automaton with terminal function ω, let $r>0$ be the maximum (positive) rank of states, and let $Q=P \cup R$, where R is the set of states with rank r. To be complete, we need the additional hypothesis that if a state q has a positive rank (resp. a negative rank) then $\omega(q) \subset A^* \times 1$ (resp. $\omega(q) \subset 1 \times B^*$). This hypothesis holds trivially for a classical automaton $\mathscr{A}=(Q, A^* \times B^*, E, I, T)$ for which $\omega(q)=1_{A^* \times B^*}$ if $q \in T$ and $\omega(q)=\emptyset$ otherwise.

Let $\mathscr{B}=(Q', A^* \times B^*, E', I, \omega')$ be the automaton defined as follows.

(i) $Q'=P \cup (R \times A)$, that is every state of maximum rank is replaced by a set of states indexed by the input alphabet; this index or second component indicates a letter that has not been read yet on the input tape, but guessed. The morphism $\varphi: \mathscr{B} \to \mathscr{A}$ is then defined by $\varphi(p)=p$ if $p \in P$ and $\varphi(r, a)=r$ if $r \in R$ and $a \in A$.

For the definition of the edges of $\mathscr{B}$, the overall idea is that the label of an edge arriving in a state (r, a) is equal to the label of the corresponding edge arriving in r multiplied on the right by $(a^{-1}, 1)$ and that the label of an edge exiting from a state (r, a) is equal to the label of the corresponding edge exiting from r multiplied on the left by $(a, 1)$. By abuse, we denote by $E \cap (S \times T)$ the set of edges of E with origin in S and end in T. It then holds that

(ii)

$$E' \cap (P \times P) = E \cap (P \times P), \tag{10}$$

that is, the edges with origin and end in P are the same in $\mathscr{B}$ and in $\mathscr{A}$.

(iii)

$$(p, (f, g), r) \in E \cap (P \times R) \;\Rightarrow\; \forall a \in A \; (p, (f, g)(a^{-1}, 1), (r, a)) \in E'. \tag{11}$$

This expression, although correct, might be misleading. Since $p \notin R$ and $r \in R$ the rank of r is strictly greater than the rank of p. The balance of (f, g) is precisely equal to the difference of ranks and is, thus, strictly positive; this implies that the length of f is $\geqslant 1$: $f = f'a$ for a certain letter a in A.

Now (11) implies that there exists one edge with origin p, with end (r, a) and with label (f', g). And that there exists no edge with origin p and with end (r, b) with $b \neq a$ for $(f, g)(b^{-1}, 1) = \emptyset$ if $b \neq a$. Thus, (10) and (11) imply that φ is a bijection between $\mathrm{Out}_{\mathscr{B}}(p)$ and $\mathrm{Out}_{\mathscr{A}}(p)$ for every p in P.

(iv)

$$(s, (f, g), q) \in E \cap (R \times P) \;\Rightarrow\; \forall b \in A \; ((s, b), (b, 1)(f, g), q) \in E'. \tag{12}$$

(v)

$$(s, (f, g), r) \in E \cap (R \times R) \;\Rightarrow\; \forall a, b \in A \; ((s, b), (b, 1)(f, g)(a^{-1}, 1), (r, a)) \in E'. \tag{13}$$

It should be noted first that since s and r have the same (maximal) rank the balance of (f, g) is zero, i.e. $|f| = |g|$.

It should be noted then that, as in (iii), for every (s, b) in $R \times A$ and every $(s, (f, g), r)$ in E, there exists only one a in A such that there exists a corresponding edge in $\mathscr{B}$ between (s, b) and (r, a):

if $|f| \geqslant 1$ then $f = f'a$ and $((s, b), (bf', g), (r, a)) \in E'$,

if $|f| = 0$ – this case might happen in the course of the iterative application of this third step because of (iii), even if one starts with a proper automaton $\mathscr{A}$ – then $|g| = 0$ and $((s, b), (1, 1), (r, b)) \in E'$.

Thus, (12) and (13) imply that φ is a bijection between $\mathrm{Out}_{\mathscr{B}}((r, a))$ and $\mathrm{Out}_{\mathscr{A}}(r)$ for every r in R and every a in A.

(vi) the definition of $\mathscr{B}$ will be complete with the one of ω':

$$\omega'(p) = \omega(p) \quad \text{if } p \in R$$

$$\omega'((r, a)) = (a, 1)\omega(r) \quad \text{if } r \in R,\ a \in A.$$

It is then a very easy verification that $\mathscr{B}$ is equivalent to $\mathscr{A}$.

A.3. Two constructions on automata related to their infinite behavior

In this section, we essentially prove two propositions that can be established for general rational ω-relations by means of Büchi's Theorem but that require direct constructions on the automata for *deterministic* relations. They both are based on the construction of a covering for the description of which it is convenient to take some conventions.

Remark A.2. If $\mathscr{B}$ is a covering of $\mathscr{A}$ (resp. an immersion in $\mathscr{A}$) then $\mathscr{B}$ is deterministic if and only if (resp. if) $\mathscr{A}$ is deterministic.

A.3.1. Notations for covering

Let $C=(A\times 1)\cup(1\times B)$ be the minimal set of generators of $A^*\times B^*$. Any 2-automaton over $A^*\times B^*$ is equivalent to a 2-automaton the edges of which are labelled with elements of C.

Let $\mathscr{A}=(Q,C,E,I,T)$ be such an automaton. A covering $\mathscr{B}$ of $\mathscr{A}$ will have $Q\times\{0,1,\ldots,n\}$ as set of states; an element (p,i) in $Q\times\{0,1,\ldots,n\}$ will be denoted by p^i and the morphism $\varphi:\mathscr{B}\to\mathscr{A}$ will be defined by $\varphi(p^i)=p$ for all p in Q and all i in $\{0,1,\ldots,n\}$. Let S be a subset of Q; then S^i denotes the subset $S\times\{i\}$.

If $\varphi:\mathscr{B}\to\mathscr{A}$ is a covering then for every p and q in Q, every i in $\{0,1,\ldots,n\}$, and every x in C, there exists a unique j in $\{0,1,\ldots,n\}$ such that (p^i,x,q^j) is an edge in $\mathscr{B}$ if and only if there exists an edge (p,x,q) in $\mathscr{A}$: the index j is a *function* of p,q,x and i.

A.3.2. Selection of the admissible infinite computations

In view of Remark A.2, Proposition 8.1 is the immediate consequence of the following proposition.

Proposition A.3. *Given a 2-automaton $\mathscr{A}$, it is possible to build effectively an immersion $\mathscr{B}$ of $\mathscr{A}$ such that*

(i) *every infinite successful computation of $\mathscr{B}$ is admissible,*

(ii) *the successful infinite computations of $\mathscr{B}$ map onto the admissible successful infinite computations of $\mathscr{A}$.*

Proof. Let $\mathscr{A}=(Q,C,E,I,T)$ be a deterministic 2-automaton (since we are interested in infinite computations, we do not need any endmarker). Let $\mathscr{C}=(Q\times\{0,1,2\}, C, E', I^0, T\times\{0,1,2\})$ be the covering of $\mathscr{A}$ determined by the following

$$(p^i,x,q^j)\in E' \Leftrightarrow (p,x,q)\in E,$$

with

$$i=0 \Rightarrow \begin{cases} j=0 \text{ if } p\notin T\ \forall x\in C,\\ j=1 \text{ if } p\in T \text{ and } x\in A\times 1,\\ j=2 \text{ if } p\in T \text{ and } x\in 1\times B,\end{cases}$$

$$i=1 \Rightarrow \begin{cases} j=1 & \text{if } x\in A\times 1,\\ j=0 & \text{if } x\in 1\times B,\end{cases}$$

$$i=2 \Rightarrow \begin{cases} j=2 & \text{if } x\in 1\times B,\\ j=0 & \text{if } x\in A\times 1.\end{cases}$$

There is a one-to-one correspondence between the (successful) infinite computations of $\mathscr{A}$ and those of $\mathscr{C}$.

Let now $\mathscr{B}$ be a subautomaton of $\mathscr{C}$: $\mathscr{B}=(Q\times\{0,1,2\}, C, E', I^0, T^0)$. $\mathscr{B}$ has the same edges as $\mathscr{C}$; it has only fewer terminal states.

From the definition of the function j above, it is clear that on any infinite computation of $\mathscr{B}$ and between two terminal states at least one letter has been read on each of the two tapes: every successful infinite computation of $\mathscr{B}$ is admissible. Conversely, every admissible successful infinite computation of $\mathscr{C}$ (and, thus, corresponding to such a computation in $\mathscr{A}$) goes infinitely often through a terminal state with index 0; it is, thus, a successful infinite computation of $\mathscr{B}$. □

A.3.3. Normalization of 2-automata with endmarker

In order to establish Proposition 8.14, we first prove a result along the same line as Proposition A.3.

Proposition A.4. *Given a 2-automaton $\mathscr{A}$ over $A^*\times B^*$ and with endmarker \$, it is possible to build effectively two immersions $\mathscr{B}_1$ and $\mathscr{B}_2$ of $\mathscr{A}$ such that*

(i) $|\mathscr{B}_1|=|\mathscr{A}|\cap(A^*\$\times B^*\$)$,

(ii) $\|\mathscr{B}_2\|=\|\mathscr{A}\|\cap(A^*\$^\omega\times B^*\$^\omega)$.

Proof. We first build a covering $\mathscr{C}$ of $\mathscr{A}$, of which $\mathscr{B}_1$ and $\mathscr{B}_2$ will be subautomata. Let $C_{\$,\$}$ denote the set $C_{\$,\$}=((A\cup\$)\times 1)\cup(1\times(B\cup\$))$ and let $\mathscr{A}=(Q, C_{\$,\$}, E, I, T)$ be a 2-automaton with endmarker. Let

$$\mathscr{C}=(Q\times\{0,1,2,3\}, C_{\$,\$}, E', I^0, T\times\{0,1,2,3\})$$

be the covering of $\mathscr{A}$ where the states with index 1, 2, or 3 characterize, respectively, the states reached after the reading of the endmarker on the first tape, on the second tape and on both tapes. More precisely, we have

$$(p^i, x, q^j)\in E' \Leftrightarrow (p, x, q)\in E$$

and

$$i=0 \Rightarrow \begin{cases} j=0 & \text{if } x\in(A\times 1)\cup(1\times B),\\ j=1 & \text{if } x=(\$,1),\\ j=2 & \text{if } x=(1,\$),\end{cases}$$

$$i=1 \Rightarrow \begin{cases} j=1 & \text{if } x\in((A\cup\$)\times 1)\cup(1\times B),\\ j=3 & \text{if } x=(1,\$),\end{cases}$$

$$i=2 \Rightarrow \begin{cases} j=2 & \text{if } x\in(A\times 1)\cup(1\times(B\cup\$)),\\ j=3 & \text{if } x=(\$,1),\end{cases}$$

$$i=3 \Rightarrow j=3 \quad \forall x\in C_{\$,\$}.$$

Let $\mathscr{B}_1 = (Q \times \{0, 1, 2, 3\}, C_{\$,\$}, E'', I^0, T^3)$ be the subautomaton of $\mathscr{C}$ obtained by keeping as final states only those with index 3 and by keeping the following edges:

- all edges with origin in Q^0;
- of the edges with origin in Q^1, only those with label in $1 \times (B \cup \$)$;
- of the edges with origin in Q^2, only those with label in $(A \cup \$) \times 1$;
- no edges with origin in Q^3.

Let $\mathscr{B}_2 = (Q \times \{0, 1, 2, 3\}, C_{\$,\$}, E''', I^0, T^3)$ be the subautomaton of $\mathscr{C}$ obtained by keeping the same terminal states as for $\mathscr{B}_1$ and by keeping the same edges as those of $\mathscr{B}_1$ plus all those with label $(\$, 1)$ or $(1, \$)$, that is:

- all edges with origin in Q^0;
- of the edges with origin in Q^1 (in Q^2), only those with label in $\{(\$, 1)\} \cup (1 \times (B \cup \$))$ (with label in $((A \cup \$) \times 1) \cup \{(1, \$)\}$);
- of the edges with origin in Q^3, only those with label $(\$, 1)$ or $(1, \$)$.

It is an easy verification to check that $\mathscr{B}_1$ and $\mathscr{B}_2$ fulfill the conditions of the proposition. □

A.3.4. ω-completion of deterministic rational relations

We are now ready to prove Proposition 8.14.

Proposition 8.14. *A rational relation of finite words is deterministic if and only if its ω-completion is a deterministic rational ω-relation.*

Proof. Let $\mathscr{A} = (Q, C_{\$,\$}, E, I, T)$ be a (deterministic) 2-automaton with endmarker and $\mathscr{B}_1 = (Q \times \{0, 1, 2, 3\}, C_{\$,\$}, E'', I^0, T^3)$ be the immersion built above. Let $\mathscr{D}$ be the automaton obtained from $\mathscr{B}_1$ by adding two states t_A and t_B which will be used to read indefinitely (but deterministically) the endmarker on the two tapes.

More precisely,

$$\mathscr{D} = ([Q \times \{0, 1, 2, 3\}] \cup \{t_A, t_B\}, C_{\$,\$}, E'' \cup F, I^0, \{t_A\}),$$

with the following new set of edges F,

$$F = \{(q^3, (\$, 1), t_B) \mid \forall q \in T\} \cup \{(t_B, (1, \$), t_A), (t_A, (\$, 1), t_B)\}.$$

Thus, a successful computation in $\mathscr{B}_1$, that is a computation that ends in a terminal state with index 3, is prolongated successfully to infinity (since t_A is terminal) and the other computations are not prolongated. If $\mathscr{A}$ is deterministic so is the immersion $\mathscr{B}_1$ and so is $\mathscr{D}$ by construction; in states of Q^3 and in t_A one reads only $(\$, 1)$ and in t_B one reads only $(1, \$)$.

Conversely, let $\mathscr{A} = (Q, C_{\$,\$}, E, I, T)$ be a 2-automaton with endmarker the infinite behavior of which is the ω-completion of a relation of finite words, that is $\|\mathscr{A}\| \subseteq A^*\$^\omega \times B^*\$^\omega$. Let $\mathscr{B}_2 = (Q \times \{0, 1, 2, 3\}, C_{\$,\$}, E''', I^0, T^3)$ be the immersion built above. Hence, $\|\mathscr{B}_2\| = \|\mathscr{A}\|$.

Any infinite computation in $\mathscr{B}_2$ has a unique decomposition into one of the two following forms:

(a) $$p^0 \xrightarrow{(u,v)} q^0 \xrightarrow{(\$,1)} r^1 \xrightarrow{(\$^n,w)} s^1 \xrightarrow{(1,\$)} t^3 \xrightarrow{(\$^\omega,\$^\omega)} \cdots,$$

(b) $$p^0 \xrightarrow{(u,v)} q^0 \xrightarrow{(1,\$)} r^2 \xrightarrow{(x,\$^m)} s^2 \xrightarrow{(\$,1)} t^3 \xrightarrow{(\$^\omega,\$^\omega)} \cdots,$$

with u, x in A^*, v, w in B^*, $n, m \geqslant 0$.

The first step in the construction is to replace the (Büchi) acceptance of infinite computations [that all have $(\$^\omega, \$^\omega)$ as label] by acceptance of finite words by terminal states.

Let $\mathscr{E} = (Q \times \{0, 1, 2, 3\}, C_{\$,\$}, F, I^0, S^3)$ be the automaton where F is obtained from E''' by deleting all edges with origin in Q^3; where S^3 is the subset of Q^3 of states that are ω-coaccessible (in $\mathscr{B}_2$).

Any computation in $\mathscr{E}$ has then a unique decomposition into one of the following two forms:

(c) $$p^0 \xrightarrow{(u,v)} q^0 \xrightarrow{(\$,1)} r^1 \xrightarrow{(\$^n,w)} s^1 \xrightarrow{(1,\$)} t^3,$$

(d) $$p^0 \xrightarrow{(u,v)} q^0 \xrightarrow{(1,\$)} r^2 \xrightarrow{(x,\$^m)} s^2 \xrightarrow{(\$,1)} t^3.$$

Hence, $\mathscr{E}$ is not yet in a normalized form for a 2-automaton with endmarker [because of the word $\$^n$ (resp. $\$^m$)] as the first (resp. second) component of the label of the computation between r^1 and s^1 (resp. between r^2 and s^2).

The second step of the construction consists in erasing the endmarker and then in making the new automaton proper. Let $\mathscr{E}' = (Q \times \{0, 1, 2, 3\}, C_{\$,\$}, F', I^0, S^3)$ be the automaton where F' is obtained from F by erasing the label $(\$, 1)$ and $(1, \$)$ in the components 1 and 2 of $\mathscr{E}$, that is

(iA) replace every edge $(p^1, (\$, 1), q^1)$ in F by an edge $(p^1, (1, 1), q^1)$,

(iB) replace every edge $(p^2, (1, \$), q^2)$ in F by an edge $(p^2, (1, 1), q^2)$,

(ii) make the new automaton proper by the classical algorithm, that is:

(iiA) replace every path

$$p^1 \xrightarrow{(1,1)} p_1^1 \xrightarrow{(1,1)} p_2^1 \xrightarrow{(1,1)} \cdots \xrightarrow{(1,1)} p_k^1 \xrightarrow{(1,b)} q^1 \tag{14}$$

by an edge $(p^1, (1, b), q^1)$ and every path

$$p^1 \xrightarrow{(1,1)} p_1^1 \xrightarrow{(1,1)} p_2^1 \xrightarrow{(1,1)} \cdots \xrightarrow{(1,1)} p_k^1 \xrightarrow{(1,\$)} s^3 \tag{15}$$

by an edge $(p^1, (1, \$), s^3)$.

(iiB) replace every path

$$p^2 \xrightarrow{(1,1)} p_1^2 \xrightarrow{(1,1)} p_2^2 \xrightarrow{(1,1)} \cdots \xrightarrow{(1,1)} p_l^2 \xrightarrow{(a,1)} q^2 \tag{16}$$

by an edge $(p^2, (a, 1), q^2)$ and every path

$$p^2 \xrightarrow{(1,1)} p_1^2 \xrightarrow{(1,1)} p_2^2 \xrightarrow{(1,1)} \cdots \xrightarrow{(1,1)} p_l^2 \xrightarrow{(\$,1)} s^3 \tag{17}$$

by an edge $(p^2, (\$, 1), s^3)$.

It is easily checked that $(u\$, v\$)$ is the label of a successful computation in $\mathscr{E}'$ if and only if $(u\$^\omega, v\$^\omega)$ is the label of a successful computation in $\mathscr{A}$. If $\mathscr{A}$ is a deterministic 2-automaton, so are $\mathscr{B}_2$ and $\mathscr{E}$. The last transformation from $\mathscr{E}$ into $\mathscr{E}'$ has to be verified more carefully. If $\mathscr{A}$ is a deterministic 2-automaton then $Q = Q_A \cup Q_B$ and an edge with origin in Q_A (resp. Q_B) has a label in $(A \cup \$) \times 1$ (resp. in $1 \times (B \cup \$)$). Hence, if there exists an edge $(p^1, (\$, 1), q^1)$ in F there exists no other edge in F with origin p^1 and with label $(\$, 1)$ or with label $(a, 1)$ – because of the definition of $\mathscr{B}_2$ – or with label in $(1 \times (B \cup \$))$ – because $Q_A \cap Q_B = \emptyset$.

Therefore, a path such as (14) or (15) is completely determined by its origin p^1; the effect of operation (iiA) (resp. (iiB)) is to reproduce the edges with origin p_k^1 (resp. p_l^2) as edges with origin p^1 (resp. p^2) and since there is no other edge with origin p^1 (resp. p^2) the automaton $\mathscr{E}'$ is deterministic as well. $\square$

Appendix B: Numeration systems and synchronized rational relations

In this appendix, we shall be interested in numbers and we shall take digits as letters that is $A = \{0, 1, \ldots, n\}$. Let us denote by $\tilde{A}$ the set $\{-n, \ldots, -1, 0, 1, \ldots, n\}$.

Let (f, g) be an element of $A^* \times A^*$. Because, in this context, f and g are considered as representations of integers, they are naturally justified *on the right* and easily identified with an element of $(A \times A)^*$ by completing the shorter by the appropriate number of zeros *on the left*. The *difference-word* of (f, g), denoted by $f \ominus g$, is defined as the word of $\tilde{A}^*$ obtained by the digit-to-digit subtraction of f by g.

By abuse, we call *set of difference-words* of a relation R of A^* into itself the set $\{f \ominus g \mid (f, g) \in R\}$ and it is denoted by $\mathrm{DW}(R)$. The following property then holds.

Proposition B.1. *The set of difference-words of a right-synchronized rational relation R is rational.*

Proof. An automaton over $\tilde{A}^*$ that recognizes the set of difference-words of R is a copy of $\mathscr{A}$, where every label (a, b) is replaced by the label $a - b$. The initial function does not bring any problem. $\square$

Corollary B.2. *The set of difference-words of a rational relation with bounded length difference is rational.*

The property expressed in Proposition B.1 is not true for every rational relation, as shown in the example below.

Example B.3. Let R be the graph of the morphism $\varphi:\{0,1\}^*\to\{0,1\}^*$ that erases the letter 0. It is easy to verify that the set of difference-words of R is $\mathrm{DW}(R)=\{e\in\{-1,0,1\}^* \mid |e|_1=|e|_{-1}\}$, which is not a rational set.

On the other hand, the set of difference-words of a relation R may well be rational, without implying that R is a rational relation.

Example B.4. Let $A=\{0,1,2\}$. Let $\sigma:A^*\to A^*$ be the substitution defined by $\sigma(0)=\sigma(1)=\{2\}$, $\sigma(2)=\{0,1,2\}$, and $\rho:A^*\to A^*$ be the mirror image of words $\rho(f_0\ldots f_i)=f_i\ldots f_0$.

Consider the relation $R=\sigma+\rho$. Then $\mathrm{DW}(R)=\{-2,-1,0,1,2\}^*$. However, R is not rational, since $R\cap(\{0,1\}^*\times\{0,1\}^*)$ is the restriction of ρ to $\{0,1\}^*$ (and would be rational with R since $\{0,1\}^*\times\{0,1\}^*$ is recognizable).

In spite of these two examples and the apparent loose connection between the rationality of a relation and the one of its set of difference-words, the latter criterion has proved to be instrumental in the characterization of numeration systems with rational normalization.

Let U be a strictly increasing sequence of integers given by a *linear recurrence relation with integers coefficients*; for instance, the sequence of Fibonacci numbers. Together with a sufficiently large alphabet of digits, U yields a numeration system in which every integer can be represented, i.e. is written as a word on the alphabet of digits. For every integer N there exist an n and digits $d_0, d_1, \ldots, d_n$ such that $N=d_0u_n+d_1u_{n-1}+\cdots+d_nu_0$; N is *represented* by the word $d_0d_1\ldots d_n$. An integer may well have several representations but the classical greedy algorithm gives a distinguished one, called the *normal representation*. The function that maps any representation onto the normal equivalent one – called the *normalization* – is, thus, a relation between words. The main problem addressed in [10, 11] is the characterization of those systems U for which the normalization is a rational function or – what is equivalent – for which the equivalence of mapping of the normalization, R_U, is a rational relation. Clearly, the set Z_U of words "equal" to zero in the numeration system U is equal to $\mathrm{DW}(R_U)$. One of the striking properties of the normalization is that R_U is a relation with bounded length difference (if the characteristic polynomial of the recurrence has a dominant root).

The result that is aimed at in this framework is the equivalence between the rationality of the normalization and the rationality of Z_U, one direction of this equivalence being, thus, given by Proposition B.1. The importance of this result comes from the possibility of characterizing the rationality of Z_U by means of algebraic tools ([10, 11]).

Using infinite words, what has been done for the representation of integers in linear numeration systems can be carried over to the representations of real numbers in a noninteger base. Recovering the usual reading from left to right, the difference-word of an element (f, g) of $A^{\mathbb{N}} \times A^{\mathbb{N}}$ is then an element of $\tilde{A}^{\mathbb{N}}$, the set of difference-words of a relation R from $A^{\mathbb{N}}$ to $A^{\mathbb{N}}$ is also denoted by $\mathrm{DW}(R)$. The same construction as above on the letter-to-letter 2-automaton yields the following proposition.

Proposition B.5. *The set of difference-words of a synchronized rational relation of infinite words is a rational set of infinite words.*

Let θ be a real number > 1. Every number of $[0, 1]$ can be represented in base θ as an infinite word on a suitable alphabet. This representation need not be unique. The representation obtained from a greedy algorithm is called the θ-development. The normalization is the function ν_θ which maps any θ-representation of a real number onto its θ-development.

In this context, the problem addressed in [10] is the characterization of those θ for which the normalization is a rational function, or – as above – for which the equivalence of mapping R_θ of the normalization is a rational relation.

The set Z_θ of elements of $\tilde{A}^*$ representing 0 in base θ is equal to $\mathrm{DW}(R_\theta)$. It is shown in [10, 11] that if ν_θ is a rational function, then R_θ is synchronized and, thus, by Proposition B.5, Z_θ is a rational set. It can be noted that, conversely, if Z_θ is a rational set, then ν_θ is a rational function. By algebraic tools, it has been proved that Z_θ is a rational set for any choice of alphabet of digits A if and only if θ is a Pisot number [11, 1].

Acknowledgments

We are pleased to thank Max Dauchet, Robert Gilman, Michel Latteux, Maryse Pelletier and Wolfgang Thomas for discussions on a first version of this paper, most of which having been made possible by the ASMICS Working Group. We also thank an anonymous referee for simplifying the proof of Proposition 5.3.

References

[1] D. Berend and Ch. Frougny, Computability by finite automata and Pisot bases, to appear.
[2] J. Berstel, *Transductions and Context-Free Languages* (Teubner, Stuttgart, 1979).
[3] L. Boasson, Cônes rationnels et familles agréables de langages – Application au langage à compteur, Thèse de 3^{e} cycle, Université Paris 7, 1971.

[4] J.R. Büchi, On a decision method in restricted second order arithmetic, in: E. Nagel et al., eds., *Logic, Methodology and Philosophy of Science* (Stanford University Press, Stanford, 1962) 1–11; reprinted in: S. Mac Lane and D. Siefkes, eds., *The Collected Works of J. Richard Büchi* (Springer, Berlin, 1990) 425–435.
[5] M. Dauchet and S. Tison, The theory of ground rewrite systems is decidable, in: *Proc. 5th IEEE Symp. LICS* (1990) 242–248.
[6] S. Eilenberg, *Automata, Languages and Machines, Vol. A* (Academic Press, New York, 1974).
[7] C.C. Elgot and J.E. Mezei, On relations defined by generalized finite automata, *IBM J. Res. Develop.* **9** (1965) 47–68.
[8] D. Epstein, J. Cannon, D. Holt, S. Levy, M. Paterson and W. Thurston, *Word Processing in Groups* (Jones and Bartlett Publishers, Boston, 1992).
[9] P.C. Fischer and A.L. Rosenberg, Multitape one-way nonwriting automata, *J. Comput. System Sci.* **2** (1968) 88–101.
[10] Ch. Frougny, Systèmes de numération linéaires et automates finis, Thèse d'Etat, Université Paris 7, LITP Report 89-69, 1989.
[11] Ch. Frougny, Representations of numbers and finite automata, *Math. Systems Theory* **25** (1992) 37–60.
[12] Ch. Frougny and J. Sakarovitch, Rational relations with bounded delay, in: *Proc. STACS 91*, Lecture Notes in Computer Science, Vol. 480 (Springer, Berlin, 1991) 50–63; also in: M. Ito, ed., *Proc. Internat. Coll. on Words, Languages and Combinatorics*, Kyoto, 1990 (World Scientific, Singapore, 1992) 145–159.
[13] F. Gire and M. Nivat, Relations rationnelles infinitaires, *Calcolo* **21** (1984) 91–125.
[14] T. Harju and J. Karhumäki, The equivalence problem of multitape finite automata, *Theoret. Comput. Sci.* **78** (1991) 347–355.
[15] J.H. Johnson, Rational equivalence relations, *Theoret. Comput. Sci.* **47** (1986) 39–60.
[16] L.H. Landweber, Decision problems for ω-automata, *Math. Systems Theory* **3** (1969) 376–384.
[17] M. Latteux and E. Timmerman, Rational ω-transductions, in: *Proc. MFCS 90*, Lecture Notes in Computer Science, Vol. 452 (Springer, Berlin, 1990) 263–277.
[18] J. Leguy, Transductions rationnelles décroissantes, *RAIRO Inform. Théor. Appl.* **15** (1981) 141–148.
[19] D. Perrin and J.E. Pin, Mots infinis, LITP Report 92–17, 1992.
[20] M.O. Rabin and D. Scott, Finite automata and their decision problems, *IBM J. Res. Develop* **3** (1959) 125–144; reprinted in: E. Moore, ed., *Sequential Machines: Selected Papers* (Addison-Wesley, Reading, MA, 1965).
[21] M.P. Schützenberger, Une variante des fonctions séquentielles, *Theoret. Comput. Sci.* **4** (1977) 47–57.
[22] J.R. Stallings, Topology of finite graphs, *Invent. Math.* **71** (1983) 551–565.
[23] W. Thomas, Automata and quantifier hierarchies, in: *Formal Properties of Finite Automata and Applications*, Lecture Notes in Computer Science, Vol. 386 (Springer, Berlin, 1988) 104–119.
[24] W. Thomas, Infinite trees and automata definable relations over ω-words, in: *Proc. STACS 90*, Lecture Notes in Computer Science, Vol. 452 (Springer, Berlin, 1990) 407–415.
[25] W. Thomas, Automata on infinite objects, in: J. van Leeuwen, ed., *Handbook of Theoretical Computer Science, Vol. B* (Elsevier, Amsterdam, 1990) 133–191.