Uniformization of rational relations

Christian Choffrut[1]
and Serge Grigorieff[2]

[1] Université Paris 7, LIAFA, 2 Pl. Jussieu, 75 251 Paris Cadex 05
[2] Université de Paris 7, 2 Pl. Jussieu, 72 251 Paris Cadex 05 France,
also L.L.A.I.C, Université de Clermont-Ferrand, France

No Title Given

No Author Given

No Institute Given

1 Introduction

Uniformizing a relation belonging to some family, consists of finding a function whose graph belongs to the family and whose domain coincides with that of the given relation. Here we are particularly concerned with the relations on finite or infinite strings that can be recognized in the traditional sense by some type of finite automaton.

Eilenberg proved in 1974 a uniformization result for rational relations on finite strings. Siefkes established in 1975 that the synchronous relations on infinite strings enjoy the uniformization property as well. Actually these results can be refined to subfamilies of rational relations on finite or infinite strings or a mixture of those: to name but the two most important, say the deterministic and the synchronous relations. Our purpose is to give a survey of all the known results of this type and to show how far they can or cannot be extended. In order to more accurately evaluate how lucky we are with dealing with strings, suffice it to say that uniformization results on trees no longer hold, see [14]. Theoretical computer science and logic have studied the subject with different tools. We think it is time to present these results in a unifying framework by bringing the two approaches together. We hope the reader will be convinced that using both the methods of theoretical computer science and those of logic helps greatly simplifying and clarifying some proofs. A good illustration is the investigation of the synchronous relations, whether on finite or infinite strings, where the language of logic spares some tedious (but of course equivalent) set constructions. This is no wonder since it allows us to use "for free" the powerful theory developed by Büchi.

Historically, the uniformization result on rational relations on finite strings can be traced back to [7], i. e., over 30 years ago. It was not stated as such, rather it was given a more precise form (technically the function by which one can uniformize a given relation is obtained as a composition of a left followed by a right "sequential" function). Since then it has been reproved with different methods, [1], [16], see also [15] for an account on the subject. Eilenberg proved it as a corollary of his "cross-section" theorem, stating intuitively that it is possible to "rationally" select a representative for each equivalence class that intersects a rational subset. This result carries over to infinite strings as well.

As previously said, most of the material here can be considered as "folklore" by some (actually non so many) researchers. There is one exception however: the proof of the uniformization property for rational relations on infinite strings seems to be new, [13]. Using a different approach, a similar result was independently obtained by D. Beauquier, J. Devolder, M. Latteux and E. Timmerman, but has not been published.

2 Preliminaries

2.1 Basics on uniformization

Let us recall a few elementary definitions in order to fix the notations. Given two sets X and Y and a (partially defined) function $f : X \to Y$, the *graph* of f is the subset $\sharp f = \{(u,v) \mid v = f(u)\}$. The *domain* $\mathrm{dom}(R)$ of a relation $R \subseteq X \times Y$ is the set of elements $x \in X$ for which there exists an element $y \in Y$ with $(x,y) \in R$. The *composition* of relations is the operation that associates with the relations $R \subseteq X \times Y$ and $S \subseteq Y \times Z$ the relation

$$R \circ S = \{(x,z) \in X \times Z \mid \text{there exists } y \in Y, \text{ with}(x,y) \in R \text{ and } (y,z) \in S\} \tag{1}$$

We compose the functions from left to right.

Now we come to the main definition of this work. Given a family $\mathcal{F}$ of relations in $X \times Y$ and $R \in \mathcal{F}$, *uniformizing* R in $\mathcal{F}$ means finding a function $f_R : X \to Y$ such that 1) $\mathrm{dom}(f_R) = \mathrm{dom}(R)$ 2) $\sharp(f_R) \subseteq R$ and 3) $\sharp(f_R) \in \mathcal{F}$. When no particular mention is given, saying that a relation belonging to a family $\mathcal{F}$ is *uniformizable*, implicitely means that it can be uniformized in $\mathcal{F}$.

2.2 Finite and infinite strings

Given a finite *alphabet* A whose elements are *symbols* or *letters*, we denote by A^* (resp. A^ω) the set of finite (resp. infinite) strings over A. As usual we denote by $A^\infty = A^* \cup A^\omega$ the set consisting of the finite and infinite strings. The empty string is denoted by ϵ regardless of which alphabet in relates to as no confusion usually arises and $|u|$ denotes the length of a string with the convention $|u| = \infty$ whenever $u \in A^\omega$.

This paper is concerned with direct products of sets of the form A^* and A^ω. One of the central tools that is used in this theory is that of *hierarchical ordering*. We recall that given a linear ordering $<$ on an alphabet A, we extend it to the free monoid A^* by posing $u <_{\mathrm{hier}} v$ if $|u| < |v|$ or if $|u| = |v|$ and there exist $w, u_1, v_1 \in A^*$ and $a, b \in A^*$, such that $u = wau_1$, $v = wbv_1$ and $a < b$ holds.

The notion of *lexicographical ordering* is more general. Consider a collection of sets E_i, where i ranges over an initial segment of the integers or over $\mathbb{N}$. Assume there exists a (possibly partial) ordering $<_i$ on each set E_i. We endow the direct product $\prod_{i \in I} E_i$ with the *lexicographical ordering* $<_{\mathrm{lex}}$ defined by $\prod_{i\in I} x_i <_{\mathrm{lex}} \prod_{i \in I} y_i$ if there exists $i \in I$ such that $x_j = y_j$ for all $j < i$ and $x_i <_i y_i$. This construction applies in particular to A^ω. Indeed, this set can be viewed as a collection of copies of A indexed by $\mathbb{N}$. Any linear ordering on A extends to a lexicographical ordering on A^ω.

As usual we will assume the set A^ω is endowed with the product topology of the discrete topology on A where the family of subsets of the form uA^ω with $u \in A^*$ form a basis of the open sets. It is a standard result that given an arbitrary lexicographical ordering on A^ω every topologically closed subset of A^ω contains its lexicographically minimal element.

3 Relations on finite strings

3.1 Rational relations on finite strings

Given an arbitrary monoid M, the least family $\mathcal{F}$ of subsets containing all finite sets and closed under set union, concatenation (i. e., X and Y are in $\mathcal{F}$ then so is $\{xy \mid x \in X, y \in Y\}$) and Kleene closure (i. e., if X is in $\mathcal{F}$ then so is $\{\epsilon\} \cup X \cup \ldots \cup X^i \ldots$) is the family of *rational* subsets and is denoted by $\mathrm{Rat}(M)$. As a particular case, given two monoids M and N, a function of M into N is *rational* if its graph is a rational subset of the product monoid $M \times N$. We refer the reader to the two handbooks [5] and [2] for basic results in this theory.

It is well-known that in the case of a direct product of free monoids $A_1^* \times \ldots \times A_n^*$, the family of rational subsets, also called *rational relations*, is precisely the family of relations recognized by finite automata. Indeed, the notion of finite automaton designed to recognize single strings, was extended in the late fifties in such a way as to operate on n-tuples of strings. The idea is to provide an automaton with as many tapes as there are components in the tuple.

More precisely assume without loss of generality that the n alphabets A_i are disjoint and set $A = \bigcup_{1 \le i \le n} A_i$. We denote by π_i the projection of A^* onto A_i^* for all $i = 1, \ldots, n$ and by π the projection onto the direct product $A_1^* \times \ldots \times A_n^*$: $\pi(w) = (\pi_1(w), \ldots, \pi_n(w))$. A finite *$n$-tape automaton* (we shall say also more simply an *automaton*) is a quadruple $\mathcal{A} = (Q, I, F, T)$ where Q is the finite set of *states*, $I \subseteq Q$ is the set of *initial* states, $F \subseteq Q$ is the set of *final* states and $T \subseteq Q \times A \times Q$ the set of *transitions*. The subset of $A_1^* \times \ldots \times A_n^*$ recognized by Aconsists of those n-tuples of strings $(\pi_1(w), \ldots, \pi_n(w))$ where w is the label of a *successful* path, i. e., a path starting in an initial state and ending in a final state (see [2, section III.6] for background on n-tape automata where they are called finite transducers).

The following is well-known, see [5, Thm IX. 2. 2.] or [2, Thm 4.1].

Theorem 1. *A relation $R \subseteq A_1^* \times \ldots \times A_n^*$ is rational if and only if there exist a rational subset $K \subseteq A^*$ such that*

$$R = \{\pi(w) \mid w \in K\}$$

There exists a deterministic version of such automata but contrarily to the free monoids, they are expressively less powerful than the non deterministic ones. Intuitively, there exists a decomposition $Q = \bigcup_{1 \leq i \leq n} Q_i$ where Q_i corresponds to the subalphabet A_i. In a state $q \in Q_i$, only transitions of letters of the subalphabet A_i are allowed and moreover a given letter defines at most one transition. Furthermore the ability to recognize the end of a component is required. More formally, we assume the alphabets A_i contain an extra symbol $\sharp_i$ (the "end-marker" of the i-th tape). We modify the π_i's by considering them as mappings of A^* into $(A_i - \sharp_i)^*$ satisfying $\pi_i(a) = a$ if $a \in A_i - \sharp_i$ and $\pi_i(a) = \epsilon$ otherwise.

An automaton is *deterministic* whenever the transitions satisfy the three conditions

$$\left.\begin{array}{l} \text{for all } (q,a,p),(q,b,r) \in T \text{ if } a \in A_i \text{ and } b \in A_j \text{ then } j = i \\ \text{for all } (q,a,p),(q,b,r) \in T \text{ if } a = b \text{ then } p = r \\ \text{for all } (q,\sharp_i,r) \in T \text{ if } w \text{ is the label of a path leaving } r \\ \qquad\qquad \text{then } w \in (A - A_i)^* \end{array}\right\} \quad (2)$$

A relation is *deterministic rational* if there exists a deterministic automaton in the above sense that recognizes it. Then the following is a paraphrase of the definition.

Proposition 1. *A relation $R \subseteq A_1^* \times \ldots \times A_n^*$ is recognized by a deterministic n-tape automaton if and only if the rational subset $K \subseteq A^*$ of Theorem 1 can be assumed to satisfy the two conditions*

$$\left.\begin{array}{l} \text{for all } u,v,w \in A^*, \text{ if } uv, uw \in K, v \in A_iA^*, w \in A_jA^* \text{ then } i = j \\ \text{for all } u \in A^*, \text{ if } u\sharp_i v \in K \text{ then } v \in (A - A_i)^* \end{array}\right\} \quad (3)$$

3.2 Synchronous relations on finite strings

Synchronous relations form an important subfamily the rational relations which enjoys nice closure properties. In particular it forms a Boolean algebra and some of its properties are decidable, whereas almost all properties of the general rational relations are undecidable (Post Correspondence Problem can be interpreted as a question on two rational relations).

Consider a fresh symbol $\sharp$ not belonging to the A_i's. With each n-tuple $(u_1, \ldots, u_n) \in \prod_{1 \leq i \leq n} A_i^*$ associate the n-tuple of strings of the same length defined as

$$(u_1, \ldots, u_n)^\sharp = (u_1\sharp^{\ell - |u_1|}, \ldots, u_n\sharp^{\ell - |u_n|}) \text{ with } \ell = \max_i |u_i| \quad (4)$$

Extending the notation to subsets $R \subseteq A_1^* \times \ldots \times A_n^*$ in the natural way, we identify $R^\sharp$ with a subset of strings over the alphabet $\Sigma = \prod_{1 \leq i \leq n} (A_i \cup$

$\{\sharp\}$). Then the relation R is *synchronous* if the subset $R^\sharp$ is recognized by a finite automaton over the alphabet Σ. It is not difficult to verify that the synchronous relations form a subfamily of the rational relations that is closed under the Boolean operations, composition of relations, direct products and projections (e. g., [7] where these relations were called FAD-relations or [8]). Finally a function $f : A_1^* \times \ldots \times A_n^* \to B_1^* \times \ldots \times B_m^*$ is *synchronous* if its graph $\sharp f$ is a synchronous relation of $A_1^* \times \ldots \times A_n^* \times B_1^* \times \ldots \times B_m^*$.

The set of synchronous relations has been logically characterized in [6]. For the reader's convenience we recall the logical language that defines it. The signature contains two symbols $<$ and E of binary predicates and a symbol T_a of unary predicate for each letter $a \in A = \bigcup_{1\leq i\leq n} A_i$. The first order language in question is defined on this signature. The individual variables belong to the disjoint union of denumerable sets X_i, for $1 \leq i \leq n$. All formulae are interpreted as follows. The universe is the union of the A_i^*'s, for $1 \leq i \leq n$, and an individual variable $x \in X_i$ is interpreted as a string in A_i^*. Now $u < v$ is true if and only if u and v belong to the same free monoid A_i^* for some $1 \leq i \leq n$ and u is a prefix of v. Furthermore, uEv is true if and only if u and v have the same length and finally $T_a(u)$ for some $a \in A$ is true if and only if the last letter of u is a. To each formula $\phi(x_1, \ldots, x_n)$ with set of free variables $x_1 \in X_{k_1}, \ldots, x_n \in X_{k_n}$ is assigned the set R of all n-tuples $(u_1, \ldots, u_n) \in A_{k_1}^* \times \ldots \times A_{k_n}^*$ such that ϕ is true when each u_i is substituted for x_i in ϕ. It is said that ϕ *defines* R or that R satisfies ϕ.

Theorem 2. *A subset $R \subseteq A_1^* \times \ldots A_n^*$ is synchronous if and only if it is defined by some formula ϕ of the above language.*

As an immediate result we get

Corollary 1. *Let $0 \leq k \leq n$ be some integer. Each synchronous relation $R \subseteq A_1^* \times \ldots \times A_n^*$ can be uniformized by some synchronous function $f : A_1^* \times \ldots \times A_k^* \to A_{k+1}^* \times \ldots \times A_n^*$.*

Proof. Observe first that the hierarchical ordering on the free monoid can be easily expressed in the logic. Also, we can express the fact that a n-tuple of strings $x_1, \ldots, x_n$ is lexicographically less than or equal to another $y_1, \ldots, y_n$. Now, let $\phi(x_1, \ldots, x_n)$ be a formula defining R. It suffices to associate with each k-tuple of $A_1^* \times \ldots \times A_k^*$ belonging to the domain of R, the (in the lexicographical ordering) least $(n-k)$-tuple of $A_{k+1}^* \times \ldots \times A_n^*$ which is associated to it. We leave it to the reader to work out the details.

3.3 Uniformization on finite strings

Intuitively, Eilenberg's cross-section theorem, [5, Thm. IX, 7. 1.] asserts that given an morphism $f : A^* \to B^*$ and a rational subset $K \subseteq A^*$, it is possible to "rationally" select a representative among all the elements of K that map onto the same element of B^*.

Theorem 3. *Let $f : A^* \to B^*$ be a morphism and let K be a rational subset of A^*. Then there exists a rational subset $L \subseteq K$ such that f maps L bijectively onto Kf.*

This result and its appoach have been widely commented, used and reproven. Traditionally, it has two major consequences: 1) each rational function f of a free monoid into another can be recognized by some "unambiguous" 2-tape automaton (each pair of strings (u, v) with $v = f(u)$ defines at most one successful path in the automaton) and 2) all rational relations of a free monoid into another are uniformizable which is precisely the result that this paper wants to extend to infinite strings, [5, Prop. IX, 8. 2].

Proposition 2. *Each rational (resp. deterministic rational) relation can be uniformized*

Proof. For rational relations this follows from Theorem 1 where the subset L of the previous theorem is substituted for K. For deterministic rational relations it suffices to observe that condition (2) still holds for all subsets of K.

Observe that the previous result cannot be extended to two or more components. Indeed, consider the following rational relation on the direct product $\{a, b\}^* \times \{a\}^* \times \{a\}^*$

$$\{(a^n b^m, a^n, a) \mid n, m \geq 0\} \cup \{(a^n b^m, a^m, \epsilon) \mid n, m \geq 0\}$$

and assume there exists a rational function $f : \{a, b\}^* \times \{a\}^* \to \{a\}^*$ that uniformizes it. Let $X_0, X_1 \subseteq \{a, b\}^* \times \{a\}^*$ be the pre-images of a and ϵ respectively and let $\mathcal{A}_0$ and $\mathcal{A}_1$ be finite automata recognizing X_0 and X_1. Denote by m the maximal number of states in these automata and by $\mu > m$ an integer which is a multiple of the number of occurrences of a (resp. b) in any simple cycle of $\mathcal{A}_0$ and $\mathcal{A}_1$ (a simple cycle is a path where initial and final states coincide and where no other state is visited more than once). Set

$$p_0 = \max\{n \in \mathbb{N} \mid (a^n b^n, a^n)f = a\} \text{ and } p_1 = \max\{n \in \mathbb{N} \mid (a^n b^n, a^n)f = \epsilon\}$$

Assume first $p_0 < \infty$ and let $N > \max\{m, p_0\}$. Consider $(a^N b^{N+\mu}, a^N) \in X_0$. By the pigeon-hole principle applied to the m first states visited by a path labelled by $(a^N b^{N+\mu}, a^N)$ there exists a cycle labelled by a pair (a^p, a^p) with $p < m$. Since μ is a multiple of p, the pair $(a^{N+\mu} b^{N+\mu}, a^{N+\mu})$ belongs to X_0, a contradiction. So we must assume that $p_0 = \infty$. A similar argument shows that $p_1 = \infty$. For some integer M greater than $m + \mu$ we have $(a^M b^M, a^M) \in X_1$. Then $(a^{M-\mu} b^M, a^{M-\mu}) \in X_0$. By the same pigeon-hole principle applied again to the m first states visited by a path labelled by $(a^{M-\mu} b^M, a^{M-\mu})$, there exists a cycle labelled by a pair (a^p, a^p) with $p < m$. Thus $(a^M b^M, a^M)$ belongs to X_0, a contradiction.

Using similar techniques we would prove that $X \cup Y \subseteq \{a\}^* \times \{a\}^* \times \{a,b\}^*$ with

$$X = \{(a^{2m}, a^p, a^m b^p) \mid m, p \geq 0\} \text{ and } Y = \{(a^{m+p}, a^m, b^m a^p) \mid m, p \geq 0\}$$

is a rational relation which cannot be uniformized by any rational function of $\{a\}^* \times \{a\}^*$ into $\{a,b\}^*$. However, when all alphabets are unary, rational relations can be uniformized for any subset of components. This follows trivially from the fact that such rational relations are defined by the logic of Presburger arithmetic, [9].

Proposition 3. *Let $1 \leq k \leq n$ be some integer and let A_i be unary alphabets for $i = 1, \ldots, n$. Each rational relation $R \subseteq A_1^* \times \ldots \times A_n^*$ can be uniformized by some rational function $f : A_1^* \times \ldots \times A_k^* \to A_{k+1}^* \times \ldots \times A_n^*$.*

4 Relations on infinite strings

Büchi generalized in [3] the notion of finite automaton in order to have it operate on infinite strings. The family of subsets of A^ω *recognized* by some Büchi automaton in this manner is denoted by Rat A^ω and is called the family of *rational subsets of infinite strings* (this is justified by the fact that this family is closed under extended "rational" operations, [5, Thm. XIV. 4. 1.]). In the same way traditional finite automata can be used to recognize relations on finite strings, Büchi automata can be used to recognize relations on infinite strings. We refer the interested reader to [11] for a thorough study of these relations. Here, we will only recall what is necessary for our purpose.

4.1 Rational relations on infinite strings

A finite n-tape automaton $\mathcal{A} = (Q, I, F, T)$ can be transformed into a *Büchi* automaton and used to recognize n-tuples of possibly infinite strings by interpreting F as a set of *repeated* states. The definitions of paragraph 3.1 carry over here naturally. An infinite path is *successful* if it starts in an initial state and visits infinitely often a repeated state. The subset of $A_1^\infty \times \ldots \times A_n^\infty$ recognized by the automaton is the set of n-tuples $\pi(w) = (\pi_1(w), \ldots, \pi_n(w))$ where $w \in A^\omega$ is the label of a successful path in the automaton (for all $i = 1, \ldots, n$ the projections π_i extend trivially from A^ω to A_i^∞).

As seen in section 3.3, the key argument for Eilenberg's uniformization result is the cross-section theorem along with what he calls the first factorization theorem (Theorem 1). For infinite strings we obtain the same result via an extension of his "second factorization theorem" which shows that all rational relations is the somposition of a synchronous relation followed by some rational substitution [5, Thm IX. 5.1.].

We recall that a morphism of B^* into A^* is *alphabetic* if it associates with every letter of B a letter of A. A *substitution* of B^* into a monoid M is a morphism of B^* into the power set of all subsets of M. In the case where M is a direct product of free monoids $A_1^* \times \ldots \times A_n^*$, the substitution can be extended from B^ω to $A_1^\infty \times \ldots \times A_n^\infty$. Hereafter we deal with substitutions into $A_1^* \times \ldots \times A_n^*$ that are *rational*, i. e., that map B^* into $\mathrm{Rat}(A_1^* \times \ldots \times A_n^*)$

Theorem 4. *[10, Prop. 2. 1.] Given a relation $R \subseteq A_1^\infty \times \ldots \times A_n^\infty$ the following conditions are equivalent.*

1) R is recognized by some Büchi automaton.

2) there exist a finite alphabet B, a rational subset $K \subseteq B^\omega$, an alphabetic morphism $\varphi : B^* \to A_1^*$ and a rational substitution $\psi : B^* \to \mathrm{Rat}(A_2^* \times \ldots \times A_n^*)$ such that $R = \varphi^{-1} \circ \cap K \circ \psi$ holds, where $\cap K$ is the retriction of the identity to the subset K.

The subsets of $A_1^\infty \times \ldots \times A_n^\infty$ recognized in this manner are called *rational relations*. It can be readily verified that for all rational relations $R \subseteq A_1^\infty \times \ldots \times A_n^\infty$ the relation $R \cap A_1^\omega \times \ldots \times A_n^\omega$ is also rational. More generally, the following holds (e. g., [11]).

Proposition 4. *For $i = 1, \ldots, n$ let $S_i = A_i^*$ or $S_i = A_i^\omega$. Let R be the relation recognized by a Büchi automaton A. Then $R \cap S_1 \times \ldots \times S_n$ is rational.*

From now on we deal with "purely infinite relations" only, i. e., with relations in $A_1^\omega \times \ldots \times A_n^\omega$. As for finite strings, the notion of deterministic automaton exists. A *deterministic* automaton satisfies the following conditions

$$\left.\begin{array}{l} \text{1) for all } (q,a,p),(q,b,r) \in T \text{ if } a \in A_i \text{ and } b \in A_j \text{ then } j = i \\ \text{2) for all } (q,a,p),(q,b,r) \in T \text{ if } a = b \text{ then } p = r \end{array}\right\} \quad (5)$$

A relation $R \subseteq A_1^\omega \times \ldots \times A_n^\omega$ is *deterministic* if there exists a deterministic automaton that recognizes it. It is *synchronous* (resp. *deterministic synchronous*) if viewed as a subset of $(A_1 \times \ldots\ldots \times A_n)^\omega$, it is recognizable by some Büchi (resp. deterministic Büchi) automaton on the alphabet $A_1 \times \ldots \times A_n$.

4.2 Uniformization on infinite strings

The main result of this paper (Theorem 5) is based on the following property which shows that synchronous relations can be uniformized.

Proposition 5. *Let $0 < k < n$ be some integer. Each synchronous relation $R \subseteq A_1^\omega \times \ldots \times A_n^\omega$ can be uniformized by some synchronous function $f : A_1^\omega \times \ldots \times A_k^\omega \to A_{k+1}^\omega \times \ldots \times A_n^\omega$.*

Proof. We first verify that it suffices to consider the case $n = 2, k = 1$. Indeed, consider two bijections $\alpha : A_1 \times \ldots \times A_k \to A$ and $\beta : A_{k+1} \times \ldots \times A_n \to B$ where A and B are new subsets. By identifying $A_1^\omega \times \ldots \times A_k^\omega$ with $(A_1 \times \ldots \times A_k)^\omega$, we may extend α to an isomorphism of $A_1^\omega \times \ldots \times A_k^\omega$ onto A^ω. Similarly we extend β to an isomorphism of $A_{k+1}^\omega \times \ldots \times A_n^\omega$ onto B^ω. The relation $\alpha^{-1} \circ R \circ \beta \subseteq A^\omega \times B^\omega$ is synchronous and can be uniformized by some function $f : A^\omega \to B^\omega$. Then the function $\alpha \circ f \circ \beta^{-1} : A_1^\omega \times \ldots \times A_k^\omega \to A_{k+1}^\omega \times \ldots \times A_n^\omega$ uniformizes R as it can be readily verified.

From now on we deal with a synchronous relation $R \subseteq A^\omega \times B^\omega$ recognized by some Büchi automaton $\mathcal{A} = (Q, I, F, T)$. It is convenient, given a pair $(u, v) \in A^\omega \times B^\omega$, to say u is the *input* and v is the *output* component. A *run* is a finite or infinite sequence of states $(q_i)_{i<n}$, $n \leq \infty$, visited in a path of the automaton, i. e., for which there exist $(u_i, v_i) \in A \times B$ such that $(q_i, u_i, v_i, q_{i+1}) \in T$ holds for all $i < n$. Without loss of generality we may enforce the following additional condition which guarantees that the output is uniquely defined by an input and a run

$$\left.\begin{array}{l} \text{for all } q, p \in Q, a \in A \text{ and } b_1, b_2 \in B \\ \text{if } (q, a, b_1, p), (q, a, b_2, p) \in T \text{ then } b_1 = b_2 \end{array}\right\} \tag{6}$$

Our proof follows the usual pattern. It consists of selecting for each input string $u \in A^\omega$ a specific image $v \in B^\omega$ satifying $(u, v) \in R$ in such a way that the selection can be performed by a finite automaton. The initial idea of Eilenberg of choosing v as the minimal string in some prescribed lexicographical ordering does not carry over to infinite strings since whatever the ordering chosen, there might not exist a minimal element associated with an input (e. g., the relation consisting of the pairs $(a^\omega, a^n b^\omega)$ and $(b^\omega, b^n a^\omega)$ for all $n \geq 0$).

In the present situation we show that to any arbitrary string $u \in A^\omega$ in the domain of the relation, we can assign a unique string $v \in B^\omega$ with $(u, v) \in R$ through a second order monadic formula. However, contrarily to Eilenberg's approach, instead of selecting the image through some of its properties we choose it via a run of the automaton. Among the runs determined by the input string, we choose that which visits repeated states earliest (hence the term "greedy ordering" see below), and whenever this does not suffice to single out one run, we will choose the minimal in the lexicographical ordering. Thus, if $\phi(u, v)$ is a monadic second order formula defining the relation R (i., e., $R = \{(u, v) \in A^\omega \times B^\omega \mid \phi(u, v) = \mathtt{true}\}$, then the uniformization is expressed by the following monadic second order formula

for all $u \in A^\omega$, $v \in B^\omega$ the three conditions hold
1) $\phi(u, v)$ is true
2) there exists a run ξ with label (u, v)
3) for all runs η with label (u, w) for some $w \neq v$, inequality $\xi < \eta$ holds

More precisely, we consider a linear ordering $<$ on Q under which the set F is an initial segment ($q \in F$ and $p < q$ implies $p \in F$) and we denote by

$<_{\text{lex}}$ the lexicographical extension of $<$ to Q^ω. Let $\top$ be a new symbol and consider the ordering on the set $F \cup \{\top\}$ which is the trace of $<$ on F and for which $\top$ is the greatest element. Extend this ordering to a lexicographical ordering on the infinite sequences on the alphabet $F \cup \{\top\}$ and denote this new ordering by $<_F$. Let finally $\pi_F : Q^\omega \to (F \cup \{\top\})^\omega$ be the substitution defined by:

$$q\pi_F = \begin{cases} q & \text{if } q \in F \\ \top & \text{otherwise} \end{cases}$$

The *greedy ordering* $<_{\text{greedy}}$ on $(Q^*F)^\omega$ is defined by setting $\eta <_{\text{greedy}} \xi$ if and only if

$$\eta\pi_F <_F \xi\pi_F \text{ or } (\ \eta\pi_F = \xi\pi_F \text{ and } \eta <_{\text{lex}} \xi\ ) \tag{7}$$

We leave it to the reader to verify that $<_{\text{greedy}}$ is indeed an ordering.

Consider an input u and let Accept_u be the set of successful runs associated with it. We assume Accept_u is non empty and we shall "construct" its $<_{\text{greedy}}$-minimal element. We start with defining the ordering $\prec$ on the set $(Q-F)^*F$ by posing $x \prec y$ if one of the following conditions holds: 1) $|x| < |y|$ or 2) $|x| = |y|$ and their last occurrences $x', y' \in F$ satisfy $x' < y'$ or 3) $|x| = |y|$ and $x' = y'$ and $x <_{lex} y$.

Let $x_1 \in (Q-F)^*F$ be the $\prec$-smallest element in $(Q-F)^*F$ which is the prefix of some run in Accept_u and let $S_1 \subseteq \text{Accept}_u$ be the non empty set of successful runs starting with x_1. Now let $x_2 \in (Q-F)^*F$ be the $\prec$-smallest string such that x_1x_2 is the prefix of some successful run in S_1 and let $S_2 \subseteq S_1$ be the non empty set of successful runs starting with x_1x_2. The process continues and defines an infinite string $x_1x_2\ldots$ which is the $<_{\text{greedy}}$-minimal element of Accept_u.

Because of condition (6) we have defined in this way a mapping $f : A^\omega \to B^\omega$ by setting $f(u) = v$ where v is the output associated with the run η. Since the greedy and the lexicographical orderings are monadic second order definable, so is the function f which completes the proof.

We are now ready to prove the result for arbitrary rational relations.

Theorem 5. *Let $0 < i < n$. Every rational relation on $A_1^\omega \times \ldots \times A_n^\omega$ can be uniformized by some rational function : $f : A_i^\omega \to \prod_{j \neq i} A_j^\omega$.*

Proof. Indeed, by Theorem 4 every relation R can be factorized into $R = \varphi^{-1} \circ \cap K \circ \psi$ where B is a finite alphabet, $K \in \text{Rat}B^*$, $\varphi : B^* \to A_1^*$ is an alphabetic morphism and $\psi : B^* \to \text{Rat}(A_2^* \times \ldots \times A_n^*)$ a rational substitution. Choose for all $b \in B$ an arbitrary element in $b\psi$ and let $\psi' : B^* \to A_2^* \times \ldots \times A_n^*$ be the resulting morphism. It suffices to show that $\varphi^{-1} \circ \cap K \circ \psi'$ can be uniformized. However the relation $\varphi^{-1} \circ \cap K$ is synchronous. By the previous theorem, it can be uniformized by a function $f : A_1^\omega \to B^\omega$. The function $f_R = f \circ \psi'$ uniformizes the relation R.

The same counter-examples of paragraph 3.3 can be adapted to infinite strings (by completing every finite string with a special symbol infinitely repeated) to show that uniformization in more than one component does not hold in general. Another interesting consequence is the fact that the cross-section property holds for infinite strings.

Corollary 2. *Let $f : A^\omega \to B^\omega$ be a morphism and let K be a rational subset of A^ω. Then there exists a rational subset $L \subseteq K$ such that f maps L bijectively onto Kf.*

Proof. Indeed, the relation $\cap Kf \circ f^{-1} \circ \cap K \subseteq B^\omega \times A^\omega$ is rational. There exists a rational function $g : B^\omega \to A^\omega$ that uniformizes it, i., e., that selects for each element $x \in Kf$ a unique element in $y \in K$ with $x = yf$. Then $L = Kfg \in \mathrm{Rat}A^\omega$ and f maps bijectively L onto Kf.

Also, since all rational subsets of infinite strings are unambiguous, [4], we have

Corollary 3. *Every rational function $f : A^\omega \to B_1^\omega \times \ldots \times B_n^\omega$ is unambiguous.*

We observe, as a easy negative result, that the family of topologically closed rational relations cannot be uniformized. Indeed, consider the closed subset on the alphabets $A = B = \{a, b\}$

$$(a^\omega \times A^\omega) \cup \{(a^n bs, bs) \mid n \geq 0, s \in A^\omega\}$$

If there would exist a function that would uniformize the relation, it would be continuous, but this is clearly impossible.

4.3 A topological interpretation

The proof of uniformization for synchronous relations can also be seen as a proof of the following result.

Proposition 6. *On every rational rational subset X of A^ω there exists a rational linear ordering $\leq$ such that every non empty $Y \subseteq X$ which is relatively closed in X has a smallest element.*

Proof. Let $\mathcal{A} = (Q, I, F, T)$ be a Büchi automaton recognizing $X \subseteq A^\omega$. We assume without loss of generality that the existence of two transitions of the form $(q, a, p), (q, b, p) \in T$ implies $a = b$. A *run* is a finite or infinite sequence of states $(q_i)_{i<n}$, $n \leq \infty$, visited in by path of the automaton, i. e., for which there exist $u_i \in A$ such that $(q_i, u_i, q_{i+1}) \in T$ holds for all $i < n$. Because of the previous condition, there exists at most one string in A^ω associated with a given run. A run is *successful* if it visits infinitely often some repeated state. For every non empty subset $Y \subseteq X$ we denote by Accept_Y the set of all successful runs associated with some element in Y.

As in Theorem 5 we can construct an infinite string $x_1 x_2 \ldots \in Q^\omega$ which is the $<_{\mathrm{greedy}}$-minimal element of Accept_Y. This string is a successful run on some input $u \in X$. Assume Y is of the form $Y = X \cap Z$ for some subset $Z \subseteq A^\omega$. Then u belongs to the adherence of Z. If Y is relatively closed, i. e., if we may assume furthermore that Z is closed, then u belongs to Y.

Now define a linear ordering $\prec$ on X as follows : $v \prec w$ if and only if the $<_{\mathrm{greedy}}$-minimal element of Accept_v is smaller than the $<_{\mathrm{greedy}}$-minimal element of Accept_w. If $Y \subseteq X$ is relatively closed in X then the string $u \in Y$ obtained as explained above is clearly the $\prec$-smallest element of Y.

This result can be viewed as the automaton version of a general topological result which states that on every Borel subset X of A^ω (in fact, also on every analytical subset) there exists a linear ordering such that every non empty relatively closed subset $Y \subseteq X$ has a smallest element. The proof of this result also uses an auxiliary greater topological set. A classical result (see [12, Thm 37. 1.]) states that X is a continuous image of the Baire space ω^ω hence the projection of some closed $R \subseteq A_1^\omega \times \omega^\omega$. The lexicographical ordering on $A_1^\omega \times \omega^\omega$ is a linear ordering such that every non empty closed set has a smallest element. It induces on X the wanted ordering defined as follows : $x \prec y$ iff the smallest element of $R \cap (\{x\} \times \omega^\omega)$ is smaller than the smallest element of $R \cap (\{y\} \times \omega^\omega)$.

Observe that the Baire space ω^ω cannot be replaced by any space B^ω with B finite since continuous images of compact spaces are compact.

In fact, in the proof of the above Proposition, the Baire space does occur implicitely in the definition of the greedy ordering. This can be seen as follows. There is a natural injection from the space of successful runs into the product $\omega^\omega \times F^\omega \times Q^\omega$ which maps a successful run onto the triple consisting of the sequence of positions of states in F, the sequence of successive states in F and the sequence of successive states. The greedy ordering on successful runs then corresponds to the lexicographic product of lexicographic orderings on the components.

References

1. A. Arnold and M. Latteux. A new proof of two theorems about rational transductions. *Theoretical Computer Science*, 8:261–263, 1979.
2. J. Berstel. *Transductions and context-free languages.* B. G. Teubner, 1979.
3. J.R. Büchi. Weak Second-Order Logic and Finite Automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.
4. O. Carton. Unambiguous Büchi automata, unpublished.
5. S. Eilenberg. *Automata, Languages and Machines*, volume A. Academic Press, 1974.

6. S. Eilenberg, C.C. Elgot, and J.C. Shepherdson. Sets recognized by n-tape automata. *Journal of Algebra*, 3:447–464, 1969.
7. C. C. Elgot and J. E. Mezei. On Relations Defined by Finite Automata. *IBM Journal*, 10:47–68, 1965.
8. C. Frougny and J. Sakarovitch. Synchronized rational relations of finite and infinite words. *Theoretical Computer Science*, 108:45–82, 1993.
9. S. Ginsburg and E. Spanier. Presburger formulas and languages. *Pacific Journal of Mathematics*, 16:285–296, 1966.
10. F. Gire. Two decidable problems for infinite words. *Information Processing Letters*, 22:135–140, 1986.
11. F. Gire and M. Nivat. Relations rationnelles infinitaires. *Calcolo*, XXI:91–125, 1984.
12. K. Kuratowski. *Topology*, volume 1. Academic Press, 1966.
13. S. Grigorieff. Uniformisation de relations rationnelles, unpublished.
14. S. Lifsches and S. Shelah. Uniformization and skolem functions in the class of trees. *Journal of Symbolic Logic*, 63:103–127, 1998.
15. J. Sakarovitch. A construction on finite automata that has remained hidden. *Theoretical Computer Science*, 204:205–231, 1998.
16. M. P. Schützenberger. Sur les relations rationnelles entre monoïdes libres. *Theoretical Computer Science*, 3:243–259, 1976.