Recent Results in the Theory of Rational Sets[†]

Jean Berstel
Jacques Sakarovitch

Laboratoire d'Informatique Théorique et Programmation,
Université Pierre et Marie Curie, Paris.

ABSTRACT

This paper presents a survey of recent results in the theory of rational sets in arbitrary monoids. Main topics considered here are : the so-called Kleene monoids (i.e. monoids where Kleene's theorem holds), rational functions and relations, rational sets in partially commutative monoids, and rational sets in free groups.

INTRODUCTION

Kleene's Theorem gives a combinatorial characterization of the subsets of a free monoid recognized by finite automata. It is well known that the theorem does not hold in arbitrary monoids. This fact leads to two directions of investigations. First, it is an interesting task to characterize those monoids for which Kleene's Theorem is true. Results in this topic are reported in Section I. Second, one may observe that Kleene's Theorem claims the equality of two families of subsets of a free monoid: the rational and the recognizable subsets. A systematic investigation of the properties of these two families of subsets in an arbitrary monoid was initiated in Eilenberg's treatise [18].

Rational relations, i.e. rational subsets of a direct product of free monoids, have been widely used, both in theoretical and in practical investigations. These relations, and the special case of rational functions, admit several specific representations which are useful in applications. Recent results are reported in Section II.

There is new interest in rational and recognizable subsets in special monoids that are not free, namely the partially commutative monoids. These monoids appear indeed to be a convenient structure for representing parallel processes. Although Kleene's Theorem does not hold in these monoids, combinatorial characterizations of their recognizable subsets have been given recently by Ochmanski [41]. In Section III, these and related results are given.

In the final section, rational and recognizable subsets in groups, and mainly in free groups are considered. It is quite satisfactory to see how classical notions in group theory have their counterpart in the theory of formal languages, and that some basic results can be formulated as

[†] This work has been partly supported by the Programme de Recherche Coordonnée "Mathématiques et Informatique" du Ministère de la Recherche et de la Technologie.

properties of rational or recognizable subsets.

Consider a monoid M. The family $Rat(M)$ of *rational subsets* of M is the least family of subsets of M containing the finite subsets, and closed under (finite) union, set product, and the star operation (which associates, to a subset K of M, the submonoid of M generated by K). To each of the *rational operations* (union, product and star) corresponds an *unambiguous* variation. A union is unambiguous if it is disjoint; a product is unambiguous if each element in the product can be factorized uniquely into factors in the corresponding subsets; a star is unambiguous if the submonoid is freely generated by the given set.

Using these unambiguous rational operations, there is, in each monoid M, a family $URat(M)$ of *unambiguous rational subsets*, which is the smallest family of subsets of M containing finite subsets and closed under unambiguous rational operations.

The third family of subsets considered is the family of *recognizable subsets* of M, denoted by $Rec(M)$. A subset K of M is *recognizable* if, and only if, there exists a finite monoid N, a morphism α from M into N and a subset F of N in such a way that $K = \alpha^{-1}(F)$.

Kleene's Theorem states that in a free finitely generated monoid A^*, the rational and the recognizable subsets coincide, i.e. $Rat(A^*) = Rec(A^*)$. The fact that deterministic automata exist for recognizing rational sets shows also that $Rat(A^*) = URat(A^*)$. It is well known that $Rec(M) \subset Rat(M)$ in any finitely generated monoid M. Let us mention that the stronger inclusion $Rec(M) \subset URat(M)$ holds if, and only if, $M \in URat(M)$.

For these definitions and results, and related topics, the reader may consult the treatise of S. Eilenberg. The present paper is intended to be a survey, and therefore no proofs are given.

I. KLEENE MONOIDS

Kleene's theorem is the basis for the study of rational sets of free monoids. In particular, one deduces from this result that rational sets in the free monoid form a Boolean algebra and are all unambiguous. We shall consider here the class of monoids in which Kleene's theorem holds, that is a monoid in which rational sets are all recognizable.

1. Definition and example

A monoid M is called a *Kleene monoid* if $Rat(M) = Rec(M)$. This definition implies that a Kleene monoid is finitely generated (since a monoid is recognizable in itself). Finite monoids, and free monoids over finite alphabets, are clearly Kleene monoids. It is not less clear that a group is a Kleene monoid if, and only if, it is finite (section IV of this paper will more precisely describe the relationship between rationality and recognizability in the free group) and that a finitely generated submonoid of a Kleene monoid is itself a Kleene monoid. We shall define infra (§3) a class of monoids which are all Kleene monoids.

Example 1 (Amar and Putzolu [3]): Starting from the study of a family of linear context-free languages (the Even Linear Languages), Amar and Putzolu have defined a monoid — indeed a new multiplicative structure on the set of words A^* — for which Kleene's theorem holds. Given any word w of A^*, let $\lambda(w)$ and $\rho(w)$ be respectively the left and the right factor of w both of length $\lfloor \frac{|w|}{2} \rfloor$, so that

$$w = \lambda(w) z \rho(w)$$

where $z \in A$ or $z = 1_{A^*}$. The product $\circ$ on A^* is then defined by

$$u \circ v = \lambda(v)u\rho(v)$$

It should be noted that the empty word 1_{A^*} is not an identity with respect to that operation: it is only a right-identity as is every single letter. In order to make A^* a monoid for that product $\circ$ it is necessary to adjoint an identity. It is shown in [3] that the monoid M_{AP} thus defined is a Kleene monoid.

2. Closure under free product

The study of Kleene monoids is only at its beginning. One closure property is the following.

Theorem (Reutenauer [44], Sakarovitch [46]): *The free product of two Kleene monoids is a Kleene monoid if, and only if, at least one of the two monoids has no divisors of the identity.*

(Recall that two elements p and q of a monoid M are divisors of the identity if they are different from 1_M and if $pq = 1_M$)

Example 2: The free product of $\mathbb{Z}/2\mathbb{Z}$ by itself is an infinite group and thus is not a Kleene monoid.

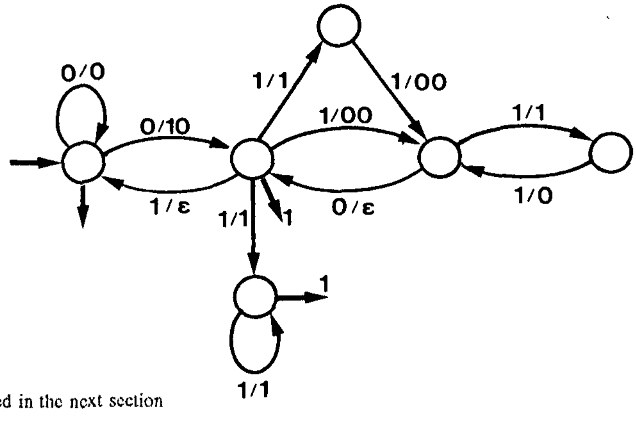
This example gives the basic idea for the proof that hypothesis of the theorem is necessary.

Example 3 (McKnight and Storey [38]): Let N be the cyclic monoid defined the relation $a = a^n$ $(n > 1)$. This monoid N has no divisors of the identity and the free product of N by itself is a Kleene monoid by the theorem above. McKnight and Storey derive this property from their study of equidivisible monoids.

3. Rational monoids

We now define a class of monoids for which it is easy to prove that they are Kleene monoids. We say that a monoid M is *rational* if there exist an alphabet A, a surjective morphism α from A^* onto M, and a rational function † β from A^* into itself which maps every word w of A^* onto a fixed representative of its class under the mapping equivalence of α.

Example 4 (the Fibonacci monoid): Let Z be the quotient of $A^* = \{0,1\}^*$ by the congruence generated by the relation $110 \equiv 001$. It is easy to see that a set of representatives for this congruence is the set $T = A^* \setminus A^*110A^*$ of words of A^* which do not contain 110 as a factor. It is remarquable that the function β which maps every word of A^* onto its representative in T is a rational function. This function may be realized for instance by the following (left) transducer:



† This notion will be defined in the next section

For a systematic exposition of this topic cf. [25] in this volume; the reader will also find there how the function β is connected to the Fibonacci numbers.

A rational monoid is a Kleene monoid. Let indeed M be a rational monoid defined as above by means of a morphism $\alpha : A^* \to M$ and of a rational function $\beta : A^* \to A^*$. If K is a rational set of M there exists a rational set L of A^* such that $\alpha(L) = K$. Hence $\alpha^{-1}(K) = \beta^{-1}\beta(L)$ is a rational set of A^* and thus K is a recognizable set of M.

Finite monoids are obviously rational monoids. From the cross-section theorem that will be considered in the next section one easily deduces that every finitely generated submonoid of a rational monoid is itself a rational monoid. The examples of Kleene monoids given above are rational monoids. For the example of McKnight and Storey, this follows from the following result:

Proposition (Sakarovitch [47]): *The free product of two rational monoids is a rational monoid if, and only if, at least one of the two monoids has no divisors of the identity.*

Up to that point the properties of Kleene monoids and of rational monoids coincide. Nevertheless examples can be constructed which show that a Kleene monoid is not necessarily a rational monoid [42].

II. RATIONAL RELATIONS AND FUNCTIONS

A *rational relation* of a monoid M into a monoid N is a rational subset of the monoid $M \times N$. The most frequent case occurs when M and N are both free monoids. In that case, $M \times N$ is not free but a very special type of partially commutative monoid. Rational relations are a widely used class of fundamental transformations, both from the theoretical and from the practical points of view (syntactic analysis, arithmetic operations, search procedures in dictionaries, decoding theory, classification of formal languages).

Elgot and Mezei [19] have shown that rational relations are realized by transducers, i.e. by finite automata with output (note that the underlying automaton may not be deterministic, and may have ε-moves). Subsequent work on rational relations has been devoted to characterizations of transducers for special classes of rational relations or functions which allow efficient implementations. Clearly, rational functions are the most important subclass. We first describe a new class of rational functions.

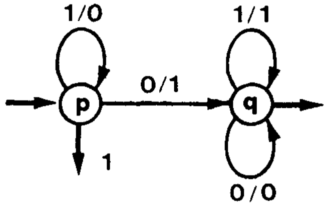
1. Plurisubsequential functions

It is well-known that it is decidable whether a given rational relation is a function (Schützenberger [50]) and moreover that an unambiguous transducer realizing the function can effectively be constructed.

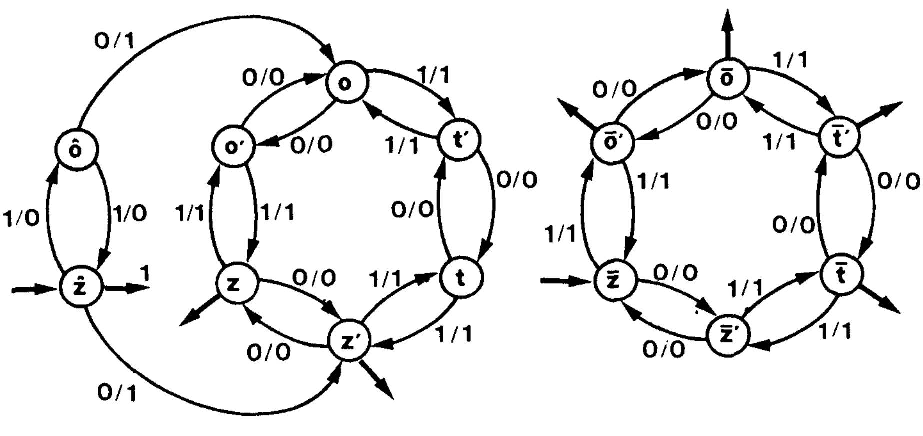
As noted by Choffrut and Schützenberger [11], the theory of rational functions was initially only a theory of finite automata with a literal output function. Elgot and Mezei then introduced the more general concept of rational functions. Although these functions have many remarkable algebraic properties, there is a annoying counterpart, since one looses the algorithmic efficiency of sequential computations. This is the reason why one has looked for larger and larger classes of rational functions which can be computed (at least partially) in a sequential manner. Thus have been defined the gsm-mappings (Ginsburg and Rose [28]), subsequential functions (Schützenberger [48], Choffrut [12]), and now plurisubsequential functions (Schützenberger and Choffrut [11]).

A *plurisubsequential function* is a finite union of subsequential functions having pairwise disjoint domains. A *subsequential function* is a function realized by a gsm equipped with an addional partial output function ρ defined on the states of the gsm. If a computation in the gsm ends in some state, then the word associated with that state by the function ρ is concatenated at the end of the output, provided ρ is defined for this state. Otherwise, it indicates that the computation is unsuccessful, and therefore that the function realized by this transducer is undefined for the given input.

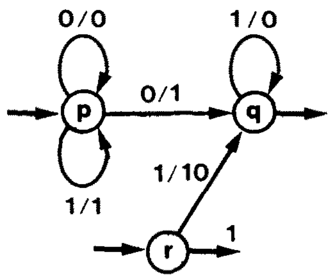
Example 1: The successor function *succ* defined over the binary expansion of nonnegative integers (with the leading bit on the right) is subsequential. It is indeed realized by the following transducer:



Example 2: Consider the function $succ_3$ from $\{0,1\}^*$ into itself defined as follows. For w a word, $succ_3(w) = succ(w)$, if w is the binary expansion (still with leading bit on the right) of an integer congruent to 0 mod. 3; otherwise $succ_3(w) = w$. The function $succ_3$ is plurisubsequential without being subsequential. It is the sum of the subsequential functions given by the following transducers:



Example 3: The successor function over the binary expansions of integers but with leading bit on the *left* (as it is usual) is a rational function, realized for instance by the following transducer:



However, we shall see as an application of the following theorem that this function is not plurisubsequential.

Theorem [11]: *Let T be an unambiguous transducer and let α be the function realized by T. The function α is plurisubsequential if, and only if, T has no branching.*

A *branching* in a transducer T is a couple (p, q) of states such that, for any integer n, there exist paths

$$p \xrightarrow{u/x} p$$

$$p \xrightarrow{u/y} q$$

in T for which the *distance* $d(x,y)$ is greater than n. As usual,

$$d(x,y) = |x| + |y| - 2|x \wedge y|$$

where $x \wedge y$ is the longest left factor of both x and y.

The pair (p, q) of states in the transducer of Example 3 is indeed a branching. To each word $u = 01^n$ there correspond two paths

$$p \xrightarrow{u/01^n} p$$

$$p \xrightarrow{u/10^n} q$$

and $d(01^n, 10^n) = 2n + 2$.

The characterization of the preceding theorem is effective. This is stated by the following

Proposition [11]: *It is decidable whether a rational function is plurisubsequential.*

2. Morphic and decreasing rational relations

It is well-known that a relation $r : A^* \to B^*$ is rational if and only if it can be factorized

$$r = \alpha \circ \cap K \circ \beta^{-1} \qquad (*)$$

where α and β are morphisms of free monoids and where $K \in Rat(C^*)$ for some alphabet C. The partial function $\cap K : C^* \to C^*$ is of course defined by $\cap K(w) = w \cap K$.

In order to have efficient computations, it is important to know how special properties of relations are reflected by the associated representations. Thus, it is known (see e.g. Eilenberg's handbook [18]) that for a length-preserving relation r (i.e. such that $v \in r(u) \Rightarrow |u| = |v|$), the morphisms α and β in the representation (*) may be chosen also to be length-preserving. This result has been strengthened as follows.

Proposition (Leguy [36]): *Let $r : A^* \to B^*$ be a rational relation which is length-decreasing* (i.e. $v \in r(u) \Rightarrow |v| \leq |u|$). *Then there is a representation*

$$r = \alpha \circ \cap K \circ \beta^{-1}$$

where α is length-decreasing and β is length-preserving.

Observe that the characterization of length-preserving rational relations is an immediate consequence of this result. Observe also that the proposition can be stated in the following way, expressing a "Fatou property": $Rat(A^* \times B^*) \cap M = Rat(M)$ where M is the submonoid of $A^* \times B^*$ given by

$$M = \{(u,v) \in A^* \times B^* \mid |u| \geq |v|\}$$

Another question investigated concerns relations which are composed only of morphisms and inverse morphisms (an arbitrary number of them may appear) but without using intersection with a rational set. Such a relation is called *morphic*. The study was motivated by a paper due to Culik II, Fich and Salomaa [15] on morphic representations of rational languages. Concerning morphic relations, let us quote the following results.

Proposition (Latteux and Leguy[35]): *Let $r : A^* \to B^*$ be a rational relation and let R be its graph. Then the following conditions are equivalent*

i) *the relation r is morphic;*
ii) *the set R is a submonoid of $A^* \times B^*$;*
iii) *there exists a factorization of r of the form*

$$r = \alpha \circ \cap(K^*) \circ \beta^{-1}$$

Proposition (Latteux and Leguy [35], Latteux and Turakainen [34]): *Any morphic relation is composed of at most four morphisms, i.e. admits representations of the form*

$$r = \alpha \circ \beta^{-1} \circ \gamma \circ \delta^{-1} \quad \textit{and} \quad r = \alpha^{-1} \circ \beta \circ \gamma^{-1} \circ \delta$$

for some morphisms α, β, γ, δ.

The proof is long and involved. It is easily verified that three morphisms do not suffice [35].

3. Rational equivalence relations and sets of representatives

Morphisms of free monoids have the following basic property, expressed in the following "Cross Section Theorem".

Theorem (Eilenberg[18]): *Let $\alpha : A^* \to B^*$ be a morphism. For any rational subset K of A^*, there exists a cross-section C of K for α which is a rational set.*

Recall that C is a *cross-section* of K for α if $\alpha(C) = \alpha(K)$, and if $\alpha|_C$ is injective.

The known proofs of Eilenberg's Theorem (Eilenberg [18], Schützenberger [49], Arnold and Latteux [6], Kobayashi [33]) all are constructive. Although these proofs do not provide a simple characterization of the computed cross-section, they suggest to consider minimal elements in each equivalence class mod. α for a given lexicographic order. In fact, one has

Proposition (Sakarovitch[45], Johnson[30]): *Let $\alpha : A^* \rightarrow B^*$ be a morphism and let K be a rational subset of A^*. For any lexicographic order on A^*, the set of minimal elements of the classes of $\alpha|_K$ is rational.*

Since a lexicographic order is not well-founded, a class $\alpha|_K$ may have no minimal elements. Thus the set

$$C = \{\min K \cap \alpha^{-1}(x) \mid x \in \alpha(K)\}$$

may not represent *all* classes of $\alpha|_K$. However, in the case where α is nonerasing, each class of α is finite and therefore has always a minimal element: consequently C is a cross-section of K for α. Observe that in the proposition, the lexicographic order cannot be replaced by the radix order, although this order is well-founded (see [45]).

The Cross Section Theorem gives in fact a property of the nuclear equivalence of morphisms of free monoids. One can look for extensions in two directions. The first consists in considering more general equivalence relations, the second concerns morphisms into monoids which are no longer free.

The first problem, on rational equivalence relations, is stated as follows: Given a rational equivalence relation R, does there exist a set of representatives of the equivalence classes which is rational. The Cross Section Theorem gives a positive answer in the case where R is the nuclear equivalence of a morphism. H. Johnson shows [30] that the answer remains positive if R is a deterministic rational equivalence relation in the sense of Fischer and Rosenberg [20]. The general case is still open [30].

Motivated by this problem, one may ask for the position of the equivalence relations among rational relations.

Proposition (Johnson [31]): *It is undecidable whether a rational relation is an equivalence relation.*

This result is similar to the well-known fact that it is undecidable whether a rational relation is recognizable. In this direction, one has

Proposition (Johnson [31]): *It is decidable whether a rational equivalence relation is recognizable.*

A second generalization of the Cross Section Theorem consists in considering the nuclear equivalence defined by morphisms into arbitrary monoids. This leads to the following definition.

A function f from A^* into a set E is *crossable* if any rational subset R of A^* has a rational cross-section for f.

As a consequence of the Cross Section Theorem, it is easily seen that if a surjective morphism $\alpha : A^* \rightarrow M$ is crossable, then any morphism from A^* into M is crossable. In this case, the monoid M itself is called *crossable*. Observe also (Choffrut, see [8]) that *any rational function from a free monoid into a crossable monoid is crossable.*

Example : A rational monoid is crossable.

Consider indeed a rational monoid M and a surjective morphism $\alpha : A^* \rightarrow M$. Let β be a rational function from A^* into itself which associates to each word $u \in A^*$ a fixed representative of the equivalence class of u modulo α. In view of the result of Choffrut mentioned above, the rational function β is crossable. Thus $\alpha = \beta \circ \alpha$ is crossable. This suffices to ensure that M itself is crossable.

There are other monoids which are crossable.

Proposition (Sakarovitch [45]): *A free group is crossable.*

On the other hand, of course not every monoid is crossable.

Example: A free commutative monoid is not crossable. Let $A = \{a,b\}$ and

$$R = a^*(ab^2)^* \cup (a^2b)^*b^*$$

There is no rational cross-section of R for the canonical morphism of A^* onto the free commutative monoid over A.

III. PARTIALLY COMMUTATIVE MONOIDS

Let A be a finite alphabet and let Θ be on symmetrical relation a A which we shall call *a commutation relation* on A. The congruence on A^* generated by the relation $ab = ba$ for all pairs (a,b) in Θ will also be denoted by Θ; the quotient A^*/Θ is *the free partially commutative monoid* generated by A with respect to the relation Θ. If Θ is empty, A^*/Θ is the free monoid A^* itself; if Θ is the universal relation (i.e. $\Theta = A{\times}A$), A^*/Θ is *the free commutative monoid* generated by A. The direct product of two free monoids that we have considered in the preceding section is another example of free partially commutative monoid.

Since long time already, the free partially commutative monoids have been considered in connection with combinatorial problems (cf. Cartier and Foata [10], Lothaire [37], Duboc [16]). More recently, words over partially commutative alphabet became of interest to computer scientists for they give a model to problems of concurrency control. In this framework, the alphabet consists in functions, and the commutation between these letters corresponds to the commutation of the composition of the corresponding functions (cf. Ullman [52] for instance). Sets of words over such partially commutative alphabets were introduced by Masurkiewicz under the name of *trace languages*, as a tool for describing the behaviour of concurrent program schemes, in the same way that classical formal languages can describe the behaviour of sequential program schemes. Three recent surveys (Masurkiewicz [39], Perrin [43], Aalbersberg and Rozenberg [2]) give rather a complete description of the subject. We shall focus here on the properties of recognizable and rational subsets of free partially commutative monoids.

1. Recognizable subsets of free partially commutative monoids

In a free partially commutative monoid, the family of recognizable sets do not coincide with the family of rational sets. Before coming to the problem of characterizing recognizable sets, let us quote an interesting closure property which is one of the oldest results on recognizable sets of free partially commutative monoids:

Proposition (Fliess [23]): *The family of recognizable sets of a free partially commutative monoid is closed under product.*

For the remaining of this section, let A be a finite alphabet, and let Θ be a commutation relation on A. We denote again by Θ the canonical (surjective) morphism from A^* onto A^*/Θ. As for the subsets of any quotient monoid, a subset T of A^*/Θ is recognizable if, and only if, $\Theta^{-1}(T)$ is a recognizable (and thus a rational) subset of A^*. If A^*/Θ is a direct product of free monoids, recognizable sets of A^*/Θ are the finite unions of direct products of rational sets of the direct components of A^*/Θ (this is a theorem due to Mezei). In the general case, recognizable sets are characterised by the following:

Theorem (Ochmanski [41]): *A subset T of A^*/Θ is recognizable if, and only if, the set $\{\min(\Theta^{-1}(t)) \mid t \in T\}$ is a rational subset of A^*.*

In this statement, as in Section II, we denote by $\min(R)$ the smallest element of the set R in a lexicographical ordering of A^*. Since $\Theta^{-1}(t)$ is finite for any t, such a smallest element always exists. Note that the fact that the lexicographic cross-section of Θ is a rational set of A^* (Anisimov and Knuth [4]) is an immediate consequence of the theorem.

The next result gives a constructive characterization of the recognizable sets that is in some sense similar to Kleene's theorem in the free monoid. Let us first define the *graph of conflicts* of an element t of A^*/Θ: the vertices of this graph are the letters which occur in t and a pair (a,b) of such vertices is an edge of this graph if (a,b) is *not* in Θ. An element t is said to be *connex* if so is its graph of conflicts. A subset T of A^*/Θ is connex if each of its elements is connex. The family $CRat(A^*/\Theta)$ is then defined to be the smallest family of subsets of A^*/Θ which contains the finite subsets and which is closed under (finite) union, product, and star restricted to connex subsets. One can then state

Theorem (Ochmanski [41]): *Let M be a free partially commutative monoid. Then $Rec(M) = CRat(M)$.*

It may be noted that this result implies earlier results giving sufficient conditions on a recognizable set T of a free partially commutative monoid for the set T^* be again recognizable (Flé and Roucairol [21], Cori and Perrin [13], Cori and Métivier [14], Métivier [40]).

2. Rational sets of a free partially commutative monoid

For further reference it is convenient to recall first the known results concerning the rational sets of commutative monoids.

Theorem A (Ginsburg and Spanier [26], Eilenberg and Schützenberger [17]): *The rational sets of a finitely generated commutative monoid form a Boolean algebra.*

Theorem B (Eilenberg and Schützenberger [17]): *The rational sets in a commutative monoid are unambiguous rational.*

Theorem C (Ginsburg and Spanier [27]): *It is decidable whether a rational set of a free commutative monoid is recognizable.*

Note that a new proof of Theorem C has recently been given by Gohon([29]). It is based on Theorem B where the original one is based on the decidability of Presburger arithmetic. An often-used example shows that the previous results will not extend without further hypothesis. The monoid $C = \{a,b\}^* \times \{c\}^*$ is a free partially commutative monoid. Let P and Q be the subsets of C defined by

$$P = (a,c)^*(b,1)^* \quad Q = (a,1)^*(b,c)^*$$

It is readily seen that

$$P \cap Q = \{(a^n b^n, c^n) \mid n \in \mathrm{N}\}$$

is not a rational subset of C and it belongs to folklore (cf. Eilenberg [18]) that $P \cup Q$ is an inherently ambiguous rational subset of C. However one has

Theorem (Aalbersberg and Welzl [1], Bertoni et al. [9], Sakarovitch [46]): *Let A be a finite alphabet and let $M = A^*/\Theta$ be a free partially commutative monoid. The following three conditions are equivalent:*

i) *the rational sets of M form a Boolean algebra;*

ii) *the rational sets of M are unambiguous rational;*

iii) *the relation Θ is transitive.*

It easily seen that the relation Θ on A is transitive if, and only if, *A^*/Θ is isomorphic to a free product of free commutative monoids.*

The above example shows that conditions (i) or (ii) imply condition (iii). Indeed, if Θ is not a transitive relation on A there exist three letters a, b, and c in A such that (a,c) and (c,b) are both in Θ while (a,b) is *not* in Θ. Thus the monoid C is isomorphic to a submonoid of A^*/Θ. Up to that isomorphism the above sets P and Q are rational subsets of C and the conclusion follows.

The converse implications may be deduced from Theorems A and B above and from the two following results ([46]).

Theorem : *Let M and N be two monoids the rational subsets of which form a Boolean algebra. Then the rational subsets of the free product $M * N$ form a Boolean algebra.*

Theorem : *Let M and N be two monoids the rational subsets of which form a Boolean algebra and are unambiguous rational. The rational subsets of the free product $M * N$ are unambiguous rational.*

As a final remark on this subject, we may observe that equality between two rational sets, as well as recognizability of a given rational set of a partially commutative monoid are both undecidable questions in the general case. However these questions become decidable in the case of a free product of free commutative monoids.

IV. FREE GROUPS

The deep connection between the structure of free groups and some chapters of formal language theory has long been recognized (cf. Chomsky-Schützenberger's theorem for context-free languages). It is also remarkable, and worth mentioning, that the most usual finiteness conditions on a subgroup of any group all have equivalent formulations in terms of formal languages theory. This can be stated as follows (for a systematic survey, see e.g. Frougny *et al.* [24])

Proposition: *Let G be a group and H a subgroup of G.*

a) *The subgroup H has finite index if, and only if, it is recognizable in G.*

b) *The subgroup H is finitely generated if, and only if, it is a rational subset of G.*

c) *The subgroup H is the normal closure of finitely many elements if, and only if, H is normal and is a context free subset of G (i.e. H is the image of a context-free subset in a surjective morphism from a free monoid onto G).*

Statement (a) in the above is indeed immediate from the definition. Statement (b) is due to Anisimov and Seiffert [5]; from their proof one can deduce a stronger statement which expresses a kind of Fatou property for groups :

Proposition: *Let H be a subgroup of a group G. A rational set of G contained in H is a rational set of H.*

Remark that this proposition does not hold in the general case, i.e. if G is replaced by any monoid M. Consider for instance the free monoid A^* with $A = \{a,b\}$ and the (recognizable) submonoid

$N = (a^*b)^*$ of A^*. This submonoid N is a rational set of A^* but is not a rational set of itself since it is not finitely generated.

Let us now turn to the free group. It is know for long time already that the rational sets of a (finitely generated) free group form a Boolean algebra (Benois [7]) and that they are all unambiguous (Fliess [22]). Since any subgroup of the free group is free and thus infinite, any finite subset of a free group is *disjunctive*, i.e. is not a union of classes for a non-trivial congruence. Hence a free group is not a Kleene monoid. Nevertheless, the following result has been proved by G. Sénizergues [51], and gives a kind of weak Kleene's theorem for the free group:

Theorem: *A rational set of a free group is either recognizable or disjunctive.*

The proof of theorem is effective and thus :

Corollary [51]: *It is decidable whether a rational set of a free group is recognizable.*

We may also note that the theorem can be generalized the following way : a group is called *virtually free* if it contains a free group of finite index. Then

Theorem [51]: *Let H be a rational subset of a virtually free group G and let N be the syntactic normal subgroup of H. Then either N or G/N is finite.*

REFERENCES

[1] Aalbersberg IJ. J. and E. Welzl, Trace languages defined by regular string languages. To appear in *RAIRO Inform. Théor.*

[2] Aalbersberg IJ. J. and G. Rozenberg, Theory of traces. Tech. Rep. 85-16, Univ. Leiden.

[3] Amar V. and G. Putzolu, Generalization of regular events. *Inform. and Control* **8** (1965), 56-63.

[4] Anisimov A. V. and D. Knuth, Inhomogeneous sorting. Unpublished manuscript.

[5] Anisimov A. V. and F. D. Seifert, Zur algebraischen Charakteristik der durch kontextfreie Sprachen definierten Gruppen, *E.I.K.* **11**, (1975), 695-702.

[6] Arnold A. and M. Latteux, A new proof of two theorems about rational transductions. *Theor. Comput. Sci.* **8** (1979), 261-263.

[7] Benois M., Parties rationnelles du groupe libre, *C.R. Acad. Sci. Paris, Ser. A*, **269** (1969), 1188-1190.

[8] Berstel J., *Transductions and Context-Free Languages*, Teubner, 1979.

[9] Bertoni A., Mauri G. and N. Sabadini, Unambiguous regular trace languages. In: *Algebra, Combinatorics, and Logic in Computer Science* (J. Demetrovics, G. Katona, and A. Salomaa, eds.), Colloquia Mathematica Soc. J. Boyai 42, North Holland (1985).

[10] Cartier P. and D. Foata, *Problèmes combinatoires de commutations et réarrangements*. Lecture Notes in Math. **85**, Springer, 1969.

[11] Choffrut Ch. and M. P. Schützenberger, Décomposition de fonctions rationnelles. In: *Proc. 3rd Symp. Theor. Aspects Comput. Sci.* (B. Monien, G. Vidal-Naquet eds.), Lect. Notes Comput. Sci. **210** (1986), 213-226.

[12] Choffrut Ch., Une caractérisation des fonctions séquentielles et des fonctions sous-séquentielles en tant que relations rationnelles. *Theor. Comput. Sci.* **5** (1977), 325-338.

[13] Cori R. and D. Perrin, Automates et commutations partielles. *RAIRO Inform. Théor.* **19** (1985), 21-32.

[14] Cori R. and Y. Métivier, Recognizable subsets of some partially abelian monoids. *Theor. Comput. Sci.* **35** (1985), 179-189.

[15] Čulik II K., E. Fich and A. Salomaa, A homomorphic characterization of regular languages. *Discrete Appl. Math.* **4** (1982), 149-152.

[16] Duboc Ch., On some equations in free partially commutative monoids. To appear in *Theor. Comput. Sci.*

[17] Eilenberg S. and M. P. Schützenberger, Rational sets in commutative monoids. *J. of Algebra* **13** (1969), 344-353.

[18] Eilenberg S., *Automata, Languages and Machines, Vol. A.* Academic Press, 1974.

[19] Elgot C. and G. Mezei, On relations defined by generalized finite automata. *IBM J. of Res. Dev.* **9** (1965), 47-65.

[20] Fischer P. and A. Rosenberg, Multitape one-way nonwriting automata. *J. Comput. and Syst. Sci.* **2** (1968), 88-101.

[21] Flé M. P. and G. Roucairol, Maximal serializability of iterated transactions. *Theor. Comput. Sci.* **38** (1985), 1-16.

[22] Fliess M., Deux applications de la représentation matricielle d'une série rationnelle non commutative. *J. of Algebra* **19** (1971), 344-353.

[23] Fliess M., Matrices de Hankel. *J. Math. pures et appl.* **53** (1974), 197-224.

[24] Frougny Ch., J. Sakarovitch and P. Schupp, Finiteness conditions on subgroups and formal language theory. Tech. Rep. Litp 85-66.

[25] Frougny Ch., Fibonacci numeration systems and rational functions. This volume.

[26] Ginsburg S. and E. H. Spanier, Bounded ALGOL-like languages. *Trans. Am. Math. Soc.* **113** (1964), 333-368.

[27] Ginsburg S. and E. H. Spanier, Semigroups, Presburger formulas, and languages. *Pacif. J. Math.* **16** (1966), 285-296.

[28] Ginsburg S. and G. Rose, A characterization of machine mappings. *Can. J. of Math.* **18** (1966), 381-388.

[29] Gohon Ph., An algorithm to decide whether a rational subset of $\mathbf{N}^k$ is recognizable. *Theor. Comput. Sci.* **41** (1986).

[30] Johnson J., Do rational equivalence relations have regular cross-section? *Proc. 12th Intern. Conf. Automata, Languages and Programming*, Lect. Notes Comput. Sci. **194** (1985), 300-309.

[31] Johnson J., A Rational Equivalence Relation Hierarchiy. Tech. Rep. University of Waterloo, November 1985.

[32] Karhumaki J. and M. Linna, A note on morphic characterization of languages. *Discrete Appl. Math.* **5** (1983), 243-246.

[33] Kobayashi K., Classification of formal languages by functional binary transductions, *nform. and Control* **15** (1969), 95-109.

[34] Latteux M., and P. Turakainen, A new normal form for compositions of morphisms and inverse morphisms. Tech. Rep. IT-66-84, Université Lille I, July 1984.

[35] Latteux M. and J. Leguy, On the composition of morphisms and inverse morphisms. *Lect. Notes Comput. Sci.* **154** (1983), 420-432.

[36] Leguy J., Transductions rationnelles décroissantes. *RAIRO Inform. Théor.* **15** (1981), 141-148.

[37] Lothaire M., *Combinatorics on Words.* Encyclop. of Math. **17**, Addison-Wesley, 1983.

[38] McKnight Jr. J. D. and A. J. Storey, Equidivisible semigroups. *J. of Algebra* **12** (1968), 24-48.

[39] Mazurkiewicz A., Traces, histories, and graphs: instances of a process monoid. *Proc. of Math. Foundations Comput. Sci. 84*, Lect. Notes Comput. Sci. **176**, (1984), 115-133.

[40] Métivier Y., Une condition suffisante de reçonaissabilité dans un monoïde partiellement commutatif. To appear in *RAIRO Inform. Théor.*

[41] Ochmanski E., Regular behaviour of concurrent systems. *Bull. European Assoc. Theor. Comput. Sci.*, October 1985.

[42] Pelletier M. and J. Sakarovitch, Easy Multiplications II. Extensions of rational monoids. In preparation.

[43] Perrin D., Words over a partially commutative alphabet. Tech. Rep. Litp 84-59.

[44] Reutenauer Ch., Sur les semi-groupes vérifiant le théorème de Kleene. *RAIRO Inform. Théor.* **19** (1985), 281-291.

[45] Sakarovitch J., Deux remarques sur un théorème de S. Eilenberg. *RAIRO Inform. Théor.* **17** (1983), 23-48.

[46] Sakarovitch J., On regular trace languages. To appear.

[47] Sakarovitch J., Easy Multiplications. I. The realm of Kleene's Theorem. To appear in *Inform. and Control.*

[48] Schützenberger M. P., Sur une variante des fonctions séquentielles. *Theor. Comput. Sci.* **4** (1977), 47-57.

[49] Schützenberger M. P., Sur les relations rationnelles entre monoïdes libres. *Theor. Comput. Sci.* **3** (1976), 243-259.

[50] Schützenberger M. P., Sur les relations fonctionnelles. In: *Automata Theory and Formal Languages, 2nd GI Conference* (H. Brakhage ed.), Lect. Notes Comput. Sci. **33** (1975), 209-213.

[51] Sénizergues G., Solution of a conjecture about rational subsets of the free group. Tech. Rep. Litp 86-3.

[52] Ullman J., *Principles of Database Systems*, Computer Science Press, 1980.