



Semigroup automata with rational initial and terminal sets

Elaine Render*, Mark Kambites

School of Mathematics, University of Manchester, Manchester M13 9PL, England, United Kingdom

ARTICLE INFO

Article history:

Received 29 November 2007

Received in revised form 4 November 2009

Accepted 7 November 2009

Communicated by D. Perrin

Keywords:

Valence automata

Semigroups

Extended automata

G-automata

Rational subsets

ABSTRACT

We consider a natural extension of the usual definition of M -automata (also known as *extended automata* or *valence automata*) which permits the automaton to utilise more of the structure of each monoid, and additionally allows us to define S -automata for S an arbitrary semigroup. In the monoid case, the resulting automata are equivalent to the *valence automata with rational target sets* which arise in the theory of regulated rewriting systems. We study these automata in the case where the register semigroup is completely simple or completely 0-simple, obtaining a complete characterisation of the classes of languages corresponding to such semigroups, in terms of their maximal subgroups. In the process, we obtain a number of results about rational subsets of Rees matrix semigroups which are likely to be of independent interest.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Recently there has been increasing interest in finite automata augmented with a memory register which stores at any moment in time an element of a given monoid M (or group G). The register is initialised with the identity element of the monoid, modified by right multiplication by monoid elements, and a word is accepted if and only if some computation reading the word returns the register to the identity with the finite state control in an accepting state. These automata are of considerable interest both to algebraists (who know them as *blind monoid automata* or M -*automata*) and to computer scientists (to whom they are often *extended finite automata* or *valence automata*). On the one hand, they provide algebraic characterizations of important language classes such as the context-free languages, the recursively enumerable languages and the blind and partially blind counter languages [6,9,12,17,18]. On the other, they give insights into computational problems in group and monoid theory [3,7,9]. They are closely connected to the theory of *rational transductions* and *relations* [1]: the languages accepted by M -automata are exactly the images under transductions of a certain language associated with M (see also [9,13,14]). They are also closely related to regulated rewriting systems, and in particular to the *valence grammars* introduced by Paun [15]: the languages accepted by M -automata are exactly the languages generated by regular valence grammars over M [5].

While M -automata with M a monoid appear at first sight to provide much more flexibility than their group counterparts, the extent to which such an automaton can fully utilise the structure of the register monoid is somewhat limited. Indeed, if the register ever contains an element of a proper ideal, then no sequence of actions of the automaton can cause it to contain the identity again; thus, the automaton has entered a “fail” state from which it can never accept a word. It follows that the automaton can make effective use of only that part of the monoid which does not lie in a proper ideal. This observation will be formalised below in Section 2, where we show that every M -automaton is equivalent to an N -automaton where N is a *simple* or *0-simple* monoid.

A natural way to circumvent this is to weaken the requirement that the identity element be the sole initial and accepting configuration of the register, and instead permit more general sets of initial and terminal configurations. Permitting more general terminal sets was suggested by Gilman [6], and the idea has also recently appeared in the study of regulated rewriting

* Corresponding author. Tel.: +44 1613068922; fax: +44 1613063220.

E-mail addresses: e.render@maths.man.ac.uk (E. Render), Mark.Kambites@manchester.ac.uk (M. Kambites).

systems, where the introduction of *valence grammars with target sets* leads naturally to a corresponding notion of *valence automata with target sets* [4,5]. An additional advantage of allowing more general initial and accepting sets is that we remove entirely the special role played by the identity, and hence the very need for an identity. The resulting automata are thus not constrained to have register monoids, but instead can use arbitrary semigroups.

If we are to retain the advantages of monoid automata as an elegant and easily manipulated way of describing important language classes, it is clearly necessary to place some kind of restriction on the class of subsets permitted for initial and terminal configurations. Possible choices include the finite subsets or the finitely generated subsemigroups, but from a computational perspective, the most natural choice seems to be the more general *rational subsets* of the semigroup. These sets, which have been the subject of intensive study by both mathematicians and computer scientists (see for example [1,11?, 19?]), are general enough to significantly increase the descriptive power of the automata, while remaining sufficiently well behaved to permit the development of a meaningful theory. Work in this area was started by Fernau and Stiebe [4] who showed that adding a rational terminal set to the definition of a G -automaton with G a group does not change the family of languages defined.

The objective of this paper, then, is to continue research into finite automata augmented with a semigroup register with rational initial and accepting subsets, and in particular the question of when such augmentation increases the accepting power of the model. In Section 2, we briefly recall some elementary definitions from the theories of formal languages, monoid automata and rational subsets. We also motivate the more general constructions which follow, by exhibiting the limitations of the capability of conventional monoid automata to make full use of the structure of a general monoid.

In Section 3 we formally define semigroup automata with rational initial and terminal sets, and obtain some foundational results about these automata and the classes of languages which they define. In Section 4 we turn our attention to Rees matrix constructions. We first study the relationship between rational subsets and Rees matrix constructions, obtaining a number of results which may be of independent interest. Some of these results are then combined with a classical theorem of Rees [16] to yield a complete description of the classes of languages accepted by S -automata with rational initial and terminal sets whenever S is a completely simple or completely 0-simple semigroup.

2. Preliminaries

Firstly, we recall some basic ideas from formal language theory. Let Σ be a finite alphabet. We denote by Σ^* the set of all words over Σ and by ϵ the empty word. Under the operation of concatenation and with the neutral element ϵ , Σ^* forms the *free monoid* on Σ . The set $\Sigma^* \setminus \{\epsilon\}$ of non-empty words forms a subsemigroup of Σ^* , called the *free semigroup* Σ^+ on Σ . A *finite automaton* over Σ^* is a finite directed graph with each edge labelled with an element of Σ , and with a distinguished *initial vertex* and a set of distinguished *terminal vertices*. A word $w \in \Sigma^*$ is *accepted* by the automaton if there exists a directed path connecting the initial vertex with some terminal vertex labelled cumulatively with w . The set of all words accepted by the automaton is denoted L or for an automaton A sometimes $L(A)$, and is called the *language* accepted by A . A language accepted by a finite automaton is called *rational* or *regular*.

More generally, if S is a semigroup then a *finite automaton over S* is a finite directed graph with each edge labelled with an element of S , and with a distinguished initial vertex and a set of distinguished terminal vertices. An element $s \in S$ is accepted by the automaton if there exists some directed path connecting the initial vertex to some terminal vertex, the product of whose edge labels is s . If S is a monoid then we admit a unique *empty path* at each vertex with label the identity element; otherwise we consider only non-empty paths. The *subset accepted* is the set of all elements accepted; a subset of S which is accepted by some finite automaton is called a *rational subset* (see for example [19]). The rational subsets of S are exactly the homomorphic images in S of regular languages.

We shall require the following result about rational subsets of groups, which is well known and easy to prove.

Proposition 2.1. *Let G be a group. If $X \subseteq G$ is rational then the subset $X^{-1} = \{x^{-1} \mid x \in X\}$ is also rational.*

Next we recall the definition of a monoid automaton. Let M be a monoid with identity 1 and let Σ be an alphabet. An *M -automaton* over Σ is a finite automaton over the direct product $M \times \Sigma^*$. We say that it accepts a word $w \in \Sigma^*$ if it accepts $(1, w)$, that is if there exists a path connecting the initial vertex to some terminal vertex labelled $(1, w)$. Intuitively, we visualize an M -automaton as a finite automaton augmented with a memory register which can store an element of M ; the register is initialised to the identity element, is modified by right multiplication by element of M , and for a word to be accepted the element present in the memory register on completion must be the identity element. We write $F_1(M)$ for the class of all languages accepted by M -automata, or equivalently for the class of languages generated by regular valence grammars over M [5].

Since an M -automaton is finite, its edge labels must lie in a finitely generated submonoid of M . For this reason, many authors restrict attention to the case that M is finitely generated. However, it is occasionally useful to consider a class of languages $F_1(M)$ where M is not finitely generated, since while this class is a union of the classes corresponding to the finitely generated submonoids of M , it may not itself be definable by a single finitely generated monoid. See, for example, [6, Theorem 6.2], for an application of this approach.

We finish this section with some brief observations concerning the extend to which an M -automaton can make use of the structure of a general monoid. In particular, we formally justify and consider the consequences of our intuitive observation, made in the introduction, that a monoid automaton uses only that part of the monoid which does not lie in a proper ideal.

Recall that an *ideal* I of a semigroup (or monoid) S is a subset I of S with the property that $S^1 I S^1 \subseteq I$, where S^1 denotes S with an identity element adjoined. To each ideal I is associated a congruence ρ_I on S such that $(s, t) \in \rho_I$ if and only if either $s, t \in I$ or $s = t$. The quotient semigroup S/ρ_I , usually just denoted S/I , is called a *Rees quotient*, and takes the form

$$S/I = \{I\} \cup \{\{x\} \mid x \in S \setminus I\}.$$

It is usual to identify $\{x\}$ with x for each $x \in S \setminus I$; the element I , which is easily seen to be a zero element in S/I , is often denoted 0 .

Proposition 2.2. *Let I be a proper ideal of a monoid M . Then $F_1(M) = F_1(M/I)$.*

Proof. Suppose $L \in F_1(M)$, and let A be an M -automaton accepting L . First notice that any path containing an edge of the form (x, w) with $x \in I$ will itself have label with first component in I ; in particular, since I is a proper ideal, $1 \notin I$ and such a path cannot be an accepting path. It follows that we may remove any such edges without changing the language accepted, and so that we may assume without loss of generality that A has no such edges. Now for any $x_1, \dots, x_n \in M \setminus I$, it follows from the definition of M/I that $x_1 \dots x_n = 1$ in M if and only if $\{x_1\} \dots \{x_n\} = \{1\}$ in M/I . Now if we let B be the (M/I) -automaton obtained from A by replacing edge labels of the form (x, w) with $(\{x\}, w)$, it follows easily that A has a path from the initial vertex to a terminal vertex labelled $(1, w)$ if and only if B has a path from the initial vertex to a terminal vertex labelled $(\{1\}, w)$. Thus, B accepts the language L and $L \in F_1(M/I)$.

Conversely, if $L \in F_1(M/I)$ then L is accepted by some (M/I) -automaton; by an argument akin to that above, we may assume without loss of generality that B has no edges labelled by the zero element I . We now obtain from B a new M -automaton A by replacing edges labels of the form $(\{x\}, w)$ with (x, w) , and argue as above to show that A accepts exactly the language L , so that $L \in F_1(M)$. \square

A semigroup or monoid is called *simple* if it does not contain any proper ideals; similarly a semigroup with a zero element 0 is called *0-simple* if its only proper ideal is $\{0\}$.

Corollary 2.3. *For every monoid M there is a simple or 0-simple monoid N such that $F_1(M) = F_1(N)$.*

Proof. If M has no proper ideals then it is simple, so we are done. Otherwise, let I be the union of all the proper ideals of M . Then I is an ideal and, since the identity element 1 does not lie in any proper ideal, $1 \notin I$ and I is a proper ideal of M . Setting $N = M/I$, it is easily verified that N is 0-simple, and by Proposition 2.2 we have $F_1(M) = F_1(M/I) = F_1(N)$ as required. \square

3. Semigroup automata

At the end of Section 2, we saw that the extent to which traditional monoid automata can utilise the full structure of a general monoid is limited. In this section, we consider a generalisation which allows us to utilise more of the structure of arbitrary monoids, and indeed semigroups.

Let S be a semigroup and Σ a finite alphabet. We define an *S-automaton* over Σ to be a finite automaton over the direct product $S \times \Sigma^*$ together with two subsets $X_0, X_1 \subseteq S$ called the *initial set* and *terminal set* respectively. The automaton accepts a word $w \in \Sigma^*$ if there exists $x_0 \in X_0$ and $x \in S$ such that $x_0 x \in X_1$, and (x, w) labels a path from the initial vertex to a terminal vertex in the automaton. We shall be particularly interested in the case where X_0 and X_1 are rational subsets of S . For S a semigroup, we let $F_{\text{Rat}}(S)$ be the class of languages accepted by S -automata with rational initial and terminal sets.

Intuitively, an S -automaton is a non-deterministic finite automaton augmented with a register which stores an element of S . The register is (non-deterministically) initialised with an element of X_0 , and a word is accepted if there is a computation which reads it and leaves the finite state control in an accept state and the register containing a value from X_1 .

Note that if $S = M$ is a monoid then an M -automaton with initial set $\{1\}$ and rational terminal set is just an *valence automaton* over M with rational target set of the kind studied by Fernau and Stiebe [4] and the present authors [17,18]. Indeed, the following proposition says that, for M a monoid, the initial set can be taken to be $\{1\}$ without loss of generality.

Proposition 3.1. *Let M be a monoid with identity 1 , and $L \subseteq \Sigma^*$ a language. If $L \in F_{\text{Rat}}(M)$ then L is accepted by an M -automaton with initial set $\{1\}$ and a rational terminal set, that is, by an valence automaton over M with rational target set.*

Proof. Let A be an M -automaton with rational initial set $X_0 \subseteq M$ and rational terminal set $X_1 \subseteq M$ which accepts the language L . Let $L' \subseteq M \times \Sigma^*$ be the full subset accepted by A interpreted as an automaton over $M \times \Sigma^*$. Since $X_0 \subseteq M$ is rational, the set

$$X'_0 = \{(x, \epsilon) \mid x \in X_0\} \subseteq M \times \Sigma^*$$

is rational. Now let

$$K = X'_0 L' = \{(x_0 x, w) \mid x_0 \in X_0, (x, w) \in L'\}.$$

Then $w \in L$ if and only if there exists $x_0 \in X_0$ and $x \in X$ such that $(x, w) \in L'$ and $x_0 x \in X_1$. But this is true exactly if $(x', w) \in K$ for some $x' \in X_1$.

Now K is a product of two rational subsets, and hence is a rational subset. Let B be a finite automaton over $M \times \Sigma^*$ recognizing K . If we interpret B as an M -automaton with initial set $\{1\}$ and terminal set X_1 , it is immediate that B accepts exactly the language L . \square

Combining Proposition 3.1 with a result of Fernau and Stiebe [4] we obtain the following.

Theorem 3.2. *If G is a group then $F_{\text{Rat}}(G) = F_1(G)$.*

Proof. If $L \in F_{\text{Rat}}(G)$ then by Proposition 3.1, L is accepted by a G -automaton with initial set $\{1\}$ and some rational terminal set X_1 , that is, by a valence automaton with rational target set. But now by [4, Theorem 3.5], L is accepted by a finite valence automaton, that is, a G -automaton, so that L is in $F_1(G)$ as required. The converse is immediate. \square

We now turn our attention to the relationship between rational relations and semigroup automata with rational initial and terminal sets. Let Ω and Σ be finite alphabets, and consider a finite automaton over the direct product $\Omega^+ \times \Sigma^*$; it recognizes a *rational relation* $R \subseteq \Omega^+ \times \Sigma^*$. The image of a language $L \subseteq \Omega^+$ under the relation R is the set of words $y \in \Sigma^*$ such that $(x, y) \in R$ for some $x \in L$.

If $X_0, X_1 \subseteq S$ then their *left quotient* is the set

$$X_0^{-1}X_1 = \{x \in S \mid x_0x = x_1 \text{ for some } x_0 \in X_0, x_1 \in X_1\}.$$

We say that a subset $X \subseteq S$ is a *rational left quotient* if there exist rational subsets $X_0, X_1 \subseteq S$ such that $X = X_0^{-1}X_1$. Note that in a group, the rational left quotients are exactly the rational subsets, but in a general semigroup this does not hold. The following statement is a semigroup analogue of [17, Proposition 3.1], which in turn generalises a well-known observation concerning M -automata (see for example [10, Proposition 2]).

Proposition 3.3. *Let X_0 and X_1 be subsets of a semigroup S , and let $L \subseteq \Sigma^*$ be a language. Then the following are equivalent:*

- (i) L is accepted by an S -automaton with initial set X_0 and terminal set X_1 ;
- (ii) there exists a finite alphabet Ω , a morphism $\omega : \Omega^+ \rightarrow S$ and a rational relation $\rho \subseteq \Omega^+ \times \Sigma^*$ such that

$$L = (X_0^{-1}X_1)\omega^{-1}\rho.$$

If S is finitely generated then the following condition is also equivalent to those above.

- (iii) for every finite choice of generators $\omega : \Omega^+ \rightarrow S$ for S , there exists a rational relation $\rho \subseteq \Omega^+ \times \Sigma^*$ such that

$$L = (X_0^{-1}X_1)\omega^{-1}\rho.$$

Proof. To show that (i) implies (ii), suppose that L is accepted by an S -automaton A with initial set X_0 and terminal set X_1 . Choose a finite alphabet Ω and a map $\omega : \Omega^+ \rightarrow S$ such that the image $\Omega^+\omega$ contains every element of S which forms the first component of an edge label in the automaton. We now obtain from A a finite automaton B over $\Omega^+ \times \Sigma^*$ by replacing each edge label (s, x) with (ω, x) for some $\omega \in \Omega^+$ is such that $\omega\omega = s$. It is a routine exercise to verify that L is the image of $(X_0^{-1}X_1)\omega^{-1}$ under the relation accepted by B .

Conversely, suppose we are given a map $\omega : \Omega^+ \rightarrow S$ and an automaton B over $\Omega^+ \times \Sigma^*$ such that L is the image under the relation accepted by B of the language $(X_0^{-1}X_1)\omega^{-1}$. We construct from B a new automaton A over $S \times \Sigma^*$ by applying the map ω to the first component of each edge label. Now interpreting A as an S -automaton with initial set X_0 and terminal rational set X_1 , it is easily seen that A accepts the language L .

Suppose now that S is finitely generated. Clearly (iii) implies (ii). Conversely, if (i) holds then we can extend ω arbitrarily to a finite choice of generators $\omega' : (\Omega')^+ \rightarrow S$ for M , and check that we still have the desired property, so that (iii) holds. \square

Note that, unlike in the monoid case [17], we cannot conclude that $F_{\text{Rat}}(S)$ is a rational cone. This is because the composition of a rational relation in $\Omega^+ \times \Sigma^*$ with a rational transduction from Σ^* to another free monoid Γ^* need not be a rational relation in $\Omega^+ \times \Gamma^*$ (although it will be rational in $\Omega^* \times \Gamma^*$).

4. Rees matrix constructions, completely simple and completely 0-simple semigroups

In this section we apply the results of the previous sections to obtain a description of language classes $F_{\text{Rat}}(S)$ for semigroups S belonging to the important classes of completely simple and completely 0-simple semigroups.

Recall that an *idempotent* in a semigroup is any element e such that $ee = e$. A non-zero idempotent e is called *primitive* if for every non-zero idempotent f such that $ef = fe = f$ we have $e = f$. A semigroup is *completely simple* [respectively, completely 0-simple] if it is simple [0-simple] and has a primitive idempotent. For more information about completely simple and completely 0-simple semigroups, see [8].

Now let T be a semigroup, 0 be a new symbol not in T and let I, J be non-empty sets. Let $P = (P_{ji})$ be a $J \times I$ matrix with entries in $T \cup \{0\}$. We define a new semigroup with set of elements

$$(I \times T \times J) \cup \{0\}$$

and multiplication defined by

$$(i, t, j)(i', t', j') = \begin{cases} (i, tP_{ji'}t', j') & \text{if } P_{ji'} \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$(i, t, j)0 = 0(i, t, j) = 00 = 0.$$

It is simple to verify that this binary operation is associative; we call the semigroup constructed in this way a *Rees matrix semigroup with zero over T* , and denote it $M^0(T; I, J; P)$. The semigroup T is called the *base semigroup* and the matrix P the *sandwich matrix* of the construction. If P contains no zero entries then $I \times T \times J$ forms a subsemigroup of $M^0(T; I, J; P)$, called a *Rees matrix semigroup (without zero) over T* and denoted $M(T; I, J; P)$.

Rees matrix semigroups play a crucial role in much of the structural theory of semigroups. Of particular importance is the case that the base semigroup T is a group G . A Rees matrix semigroup with zero over a group is called *regular* if every row and every column of the sandwich matrix contains a non-zero entry. The importance of this construction can be seen from the following seminal result of Rees [2,16].

Theorem 4.1 (The Rees Theorem). *Let $S = M(G; I, J; P)$ [respectively, $S = M^0(G; I, J; P)$] be a [regular] Rees matrix semigroup over a group. Then S is a completely simple [respectively, completely 0-simple] semigroup. Conversely, every completely simple [completely 0-simple] semigroup is isomorphic to one constructed in this way.*

We shall need the following proposition.

Proposition 4.2. *Let $S = M(T; I, J; P)$ or $S = M^0(T; I, J; P)$ be a Rees matrix semigroup with or without zero over a semigroup T . Let $X \subseteq S$ be a rational subset and suppose $i \in I$ and $j \in J$. Then the set*

$$X_{ij} = \{g \in T \mid (i, g, j) \in X\} \subseteq T$$

is a rational subset of T .

Proof. Let A be a finite automaton over S accepting the rational subset X , with vertex set Q . Let J' be the set of all $j \in J$ such that A has an edge label with third component j ; note that J' is necessarily finite. We construct from A a new finite automaton B over T with

- vertex set $(Q \times J') \cup \{q'_0\}$ where q'_0 is a new symbol;
- start vertex q'_0 ;
- terminal vertices (q, j) such that q is a terminal vertex of A ;
- an edge from q'_0 to (q_1, j_1) labelled t_1 whenever A has an edge from the initial vertex to q_1 labelled (i, t_1, j_1) ;
- for every $j_1 \in J'$, an edge from (q_1, j_1) to (q_2, j_2) labelled $P_{j_1 j_2} t_2$ whenever A has an edge from q_1 to q_2 labelled (i_2, t_2, j_2) with $P_{j_1 j_2} \neq 0$.

Since J' is finite and A has finitely many vertices and edges, we deduce that B has finitely many vertices and edges. Now we show that the subset accepted by B is exactly X_{ij} . Let $t \in X_{ij}$. Then $(i, t, j) \in X$ labels a path in A from the initial vertex to some terminal vertex. Clearly this path cannot contain edges labelled 0, so it must have the form

$$p_0 \xrightarrow{(i_1, t_1, j_1)} p_1 \xrightarrow{(i_2, t_2, j_2)} p_2 \xrightarrow{(i_3, t_3, j_3)} \dots \xrightarrow{(i_{m-1}, t_{m-1}, j_{m-1})} p_{m-1} \xrightarrow{(i_m, t_m, j_m)} p_m$$

where p_0 is the initial vertex of A and p_m is a terminal vertex. Since the path is labelled (i, t, j) we must have

$$(i, t, j) = (i_1, t_1, j_1)(i_2, t_2, j_2) \dots (i_m, t_m, j_m)$$

so that $i_1 = i, j_m = j$. Now it follows easily from the construction of B that it has a path

$$q'_0 \xrightarrow{t_1} (p_1, j_1) \xrightarrow{P_{j_1 j_2} t_2} (p_2, j_2) \dots (p_{m-1}, j_{m-1}) \xrightarrow{P_{j_{m-1} j_m} t_m} (p_m, j),$$

where (p_m, j) is a terminal vertex of B , so that B accepts

$$t = t_1 P_{j_1 j_2} t_2 P_{j_2 j_3} \dots P_{j_{m-1} j_m} t_m.$$

Thus $X_{ij} \subseteq L(B)$.

Conversely, assume that $t \in T$ is accepted by B . Then there exists a path through B from the initial vertex to some terminal vertex labelled with t . It follows from the definition of B that this path must have the form

$$q'_0 \xrightarrow{t_1} (p_1, j_1) \xrightarrow{P_{j_1 j_2} t_2} (p_2, j_2) \dots (p_{m-1}, j_{m-1}) \xrightarrow{P_{j_{m-1} j_m} t_m} (p_m, j),$$

where p_m is a terminal vertex in A ,

$$t = t_1 P_{j_1 j_2} t_2 P_{j_2 j_3} t_3 \dots P_{j_{m-1} j_m} t_m$$

and A has a path

$$p_0 \xrightarrow{(i_1, t_1, j_1)} p_1 \xrightarrow{(i_2, t_2, j_2)} p_2 \xrightarrow{(i_3, t_3, j_3)} \dots \xrightarrow{(i_{m-1}, t_{m-1}, j_{m-1})} p_{m-1} \xrightarrow{(i_m, t_m, j)} p_m$$

where p_0 is the initial vertex of A . Hence, A accepts the element

$$\begin{aligned} (i, t_1, j_1)(i_2, t_2, j_2) \dots (i_m, t_m, j) &= (i, t_1 P_{j_1 j_2} t_2 P_{j_2 j_3} t_3 \dots P_{j_{m-1} j_m} t_m, j) \\ &= (i, t, j). \end{aligned}$$

So $(i, t, j) \in X$ and hence $t \in X_{ij}$.

So the automaton B accepts exactly the set X_{ij} , and hence X_{ij} is a rational subset of T . \square

As a corollary, we obtain a result about the intersections of rational subsets with maximal subgroups in completely simple semigroups.

Corollary 4.3. *Let H be a maximal subgroup of a completely simple or completely 0-simple semigroup S . Let X be a rational subset of S . Then $X \cap H$ is a rational subset of H .*

Proof. By the Rees theorem, we may assume that S is a Rees matrix semigroup without zero $M(G; I, J; P)$ or a regular Rees matrix semigroup with zero $S = M^0(G; I, J; P)$ over a group G . It follows easily from the definition of the Rees matrix construction that either $H = \{0\}$ or

$$H = \{(i, g, j) \mid g \in G\}$$

for some $i \in I$ and $j \in J$ with $P_{ji} \neq 0$. In the former case the result is trivial, so we assume the latter. By Proposition 4.2, the set

$$X_{ij} = \{g \in G \mid (i, g, j) \in X\} = \{g \in G \mid (i, g, j) \in H \cap X\}$$

is a rational subset of G . It follows that

$$P_{ji}X_{ij} = \{P_{ji}g \mid g \in X_{ij}\} = \{P_{ji}g \mid (i, g, j) \in X\}$$

is also a rational subset of G . Now define a map

$$\phi : G \rightarrow H, g \mapsto (i, P_{ji}^{-1}g, j)$$

where P_{ji}^{-1} is the inverse of P_{ji} in the group G . It is readily verified that ϕ is an isomorphism from G to H , and so the image

$$\begin{aligned} (P_{ji}X_{ij})\phi &= \{(i, P_{ji}^{-1}g, j) \mid g \in P_{ji}X_{ij}\} \\ &= \{(i, P_{ji}^{-1}P_{ji}g, j) \mid (i, g, j) \in X\} \\ &= \{(i, g, j) \mid (i, g, j) \in X\} \\ &= X \cap H \end{aligned}$$

is a rational subset of G , as required. \square

In a completely simple semigroup, where every element lies in a maximal subgroup, Corollary 4.3 easily yields the following complete characterisation of rational subsets.

Theorem 4.4. *The rational subsets of a completely simple semigroup are exactly the finite unions of rational subsets of maximal subgroups.*

Proof. Let S be a completely simple semigroup. If X_1, \dots, X_n are rational subsets of maximal subgroups of S then certainly they are rational subsets of S , and so is their union. Conversely, suppose X is a rational subset of S . It follows easily from the Rees theorem that X lies inside a finitely generated completely simple subsemigroup S' of S . Now S' is the union of finitely many maximal subgroups, so X is the union of its intersections with these subgroups. By Corollary 4.3 these intersections are rational, so X is a finite union of rational subsets of maximal subgroups of S' . But maximal subgroups of S' are subgroups of S , and hence lie in maximal subgroups of S . It follows that X is a finite union of rational subsets of maximal subgroups of S , as required. \square

Proposition 4.5. *Let $S = M(T; I, J; P)$ or $S = M^0(T; I, J; P)$ be a Rees matrix semigroup with or without zero over a semigroup T , and let $P' \subseteq T$ be the set of non-zero entries of the sandwich matrix P . Suppose $T = P'T$ or $T = TP'$. Then for any $i \in I, j \in J$ and rational subset X of T , the set*

$$\{(i, t, j) \mid t \in X\}$$

is a rational subset of S .

Proof. By symmetry of assumption, it suffices to consider the case in which $T = P'T$. Let A be a finite automaton over T accepting X , with vertex set Q . Let $Y \subseteq T$ be the set of edge labels in A , and for every $t \in Y$, let $j_t \in J, i_t \in I$ and $s_t \in T$ be such that $t = P_{j_t i_t} s_t$. Let $J' = \{j_t \mid t \in Y\} \cup \{j\}$. Then J' is a finite subset of J . We define a new automaton B over S with

- vertex set $(Q \times J') \cup \{q_0\}$ where q_0 is a new symbol;
- initial vertex q_0 ;
- terminal vertices (q, j) such that q is a terminal vertex of A ;
- for every edge in A from the start vertex to a vertex q labelled t , and every $j' \in J'$, an edge from q_0 to (q, j') labelled (i, t, j') ;
- for every edge in A from a vertex p to a vertex q labelled t , and every $j' \in J'$, an edge from (p, j_t) to (q, j') labelled (i_t, s_t, j') .

A routine argument, akin to that in the proof of Proposition 4.2, shows that B accepts the required subset of S . \square

Note in particular that the conditions on the sandwich matrix in the hypothesis of Proposition 4.5 are satisfied in the case of a regular Rees matrix construction over a group.

Recall that the *rational subset problem* for a finitely generated semigroup S is the algorithmic problem of deciding, given a rational subset (described as a finite automaton over a fixed generating set for S) and an element (described as a word over

the same generating set), deciding whether the latter belongs to the former. While the phrasing of the problem is dependent on the precise choice of finite generating set, the decidability or undecidability of the problem is independent of this choice [11, Corollary 3.4], so one can meaningfully say that the abstract semigroup S has decidable or undecidable rational subset problem.

Corollary 4.6. *Let $S = M(T; I, J; P)$ or $S = M^0(T; I, J; P)$ be a finitely generated Rees matrix semigroup with or without zero over a semigroup T . If T has decidable rational subset problem then S has decidable rational subset problem.*

Proof. We prove the statement for Rees matrix semigroups with zero. The result for Rees matrix constructions without zero can be obtained as an easy consequence, or proved directly using a similar method.

Let $\omega : \Omega^* \rightarrow T$ and $\sigma : \Sigma^* \rightarrow S$ be finite choices of generators for T and S respectively. For every $x \in \Sigma$ such that $x\sigma \neq 0$, suppose $x\sigma = (i_x, g_x, j_x)$ and let $w_x \in \Omega^*$ be a word with $w_x\omega = g_x$. For $j \in J$ and $i \in I$ such that $P_{ji} \neq 0$ let $w_{ji} \in \Omega^*$ be a word with $w_{ji}\omega = P_{ji}$.

Now suppose we are given a word $w = w_1 \dots w_n \in \Sigma^*$, where each $w_i \in \Sigma$, and a rational subset X of S . Clearly, we can test whether w represents 0 and, in the case that it does, whether $0 \in X$. Assume now that w does not represent 0. Then

$$w\omega = (w_1\omega) \dots (w_n\omega) = (i_{w_1}, g_{w_1} P_{j_{w_1} i_{w_2}} g_{w_2} \dots g_{w_n}, j_{w_n}).$$

Let $Y = \{t \in T \mid (i_{w_1}, t, j_{w_n}) \in X\}$, so that $w\omega \in X$ if and only if

$$(g_{w_1} P_{j_{w_1} i_{w_2}} g_{w_2} \dots g_{w_n})\sigma = g_{w_1} P_{j_{w_1} i_{w_2}} g_{w_2} \dots g_{w_n} \in Y. \quad (1)$$

Now by Proposition 4.2, Y is rational and it follows moreover from the proof that we can effectively compute an automaton for Y . By assumption, we can solve the rational subset problem for Y , so we can decide whether (1) holds, as required. \square

We now turn our attention to languages accepted by S -automata with rational initial and terminal sets, where S is a Rees matrix semigroup. We begin with a lemma which simplifies the case of Rees matrix semigroups with zero, by allowing us to restrict attention to automata for which neither the initial set nor the terminal set contains the zero element.

Lemma 4.7. *Let $S = M^0(T; I, J; P)$ be a finitely generated Rees matrix semigroup with zero over a semigroup T , and suppose P contains a non-zero entry. If $L \in F_{\text{Rat}}(S)$ then L is accepted by an S -automaton with rational initial and terminal sets neither of which contain 0.*

Proof. Suppose L is accepted by an S -automaton A with rational initial set X_0 and rational terminal set X_1 . Suppose first that $0 \in X_0$. If also $0 \in X_1$ then we have $0x \in X_1$ for all $x \in S$, so the language accepted is just the set of all words w such that (x, w) labels a path from the initial vertex to a terminal vertex of A for some $x \in S$. It follows that L is regular, from which it can be easily deduced that L is accepted by an S -automaton with rational initial and terminal set $S \setminus \{0\}$. On the other hand, if $0 \notin X_1$ then there is no $x \in S$ such that $0x \in X_1$; hence we may replace the initial set X_0 with $X_0 \setminus \{0\}$ without changing the language accepted. Thus, we may assume that $0 \notin X_0$.

Clearly we can write $L = L_0 \cup L_1$ where L_1 is accepted by a S -automaton with 0 not in the rational initial or terminal sets, and L_0 is accepted by an S -automaton with rational initial set and terminal set $\{0\}$. We claim that L_0 is regular; it will follow that L is the union of L_1 with a regular language, and so can be accepted by an S -automaton with rational initial and terminal sets not containing 0 in the terminal set.

Let $\omega : \Omega^* \rightarrow S$ be a finite choice of generators for S . For each $x \in \Omega$ such that $x\omega \neq 0$ suppose $x\omega = (i_x, g_x, j_x)$. Now let K be the set of all words representing elements of the initial set of A , and let K' be the (necessarily finite) set of all final letters of words in K . It is easily seen that the language

$$\{v \in \Omega^* \mid (wv)\omega = 0 \text{ for some } w \in K\}$$

is regular. Indeed, it consists of all words which

- contain a generator representing zero; or
- contain consecutive generators x and y with $P_{j_x i_y} = 0$; or
- start with a generator y with $P_{j_x i_y} = 0$ for some $x \in K'$

and so can be easily described by a regular expression. It now follows from Proposition 3.3 that L_0 is a rational transduction of the above regular language and hence is itself regular. This completes the proof. \square

We are now ready to prove the main theorem of this section.

Theorem 4.8. *Let S be a completely simple or completely 0-simple semigroup with maximal non-zero subgroup G . Then*

$$F_{\text{Rat}}(S) = F_{\text{Rat}}(G) = F_1(G).$$

Proof. That $F_{\text{Rat}}(G) = F_1(G)$ is Theorem 3.2, while the inclusion $F_{\text{Rat}}(G) \subseteq F_{\text{Rat}}(S)$ is immediate. Hence, we need only prove that $F_{\text{Rat}}(S) \subseteq F_{\text{Rat}}(G)$. It follows easily from the Rees theorem that every completely simple semigroup S embeds in a completely 0-simple semigroup S' with the same maximal non-zero subgroup, so that $F_{\text{Rat}}(S) \subseteq F_{\text{Rat}}(S')$. Hence, it suffices to prove the result in the case that S is completely 0-simple.

Suppose, then, that S is completely 0-simple. By the Rees theorem, we may assume that S is a regular Rees matrix semigroup $M^0(G^0; I, J; P)$ where G is a group. Suppose now that a language $L \subseteq \Sigma^*$ lies in $F_{\text{Rat}}(S)$. Let A be an S -automaton accepting L , with initial rational set $X_0 \subseteq S$ and terminal rational set $X_1 \subseteq S$. By Lemma 4.7, we may assume that $0 \notin X_0$ and $0 \notin X_1$.

Let C and D be automata over S accepting X_0 and X_1 respectively. Since C, D and A have only finitely many edges between them, we may choose finite subsets $I' \subseteq I$ and $J' \subseteq J$ such that the edge labels of C and D all lie in $I' \times G \times J'$, and the edge labels of A all lie in $(I' \times G \times J') \times \Sigma^*$.

For each $i \in I'$ and $j \in J'$, we let $X_{ij} = \{g \in G \mid (i, g, j) \in X_0\}$. By Proposition 4.2, each X_{ij} is a rational subset of G . It follows that

$$X'_{ij} = X_{ij} \times \{\epsilon\}$$

is a rational subset of $G \times \Sigma^*$; let C_{ij} be an automaton accepting X'_{ij} .

Similarly, for each $i \in I'$ and $j \in J'$ we define $Y_{ij} = \{g^{-1} \in G \mid (i, g, j) \in X_1\}$. By Propositions 4.2 and 2.1, Y_{ij} is a rational subset of G , and so

$$Y'_{ij} = Y_{ij} \times \{\epsilon\}$$

is a rational subset of $G \times \Sigma^*$; let D_{ij} be an automaton accepting Y'_{ij} .

Assume without loss of generality that the automaton A and all the automata C_{ij} and D_{ij} have disjoint vertex sets. We construct from these automata a G -automaton B with

- vertex set the union of the vertex sets of C_{ij} and D_{ij} (for $i \in I'$ and $j \in J'$) together with $I' \times Q \times J'$ where Q is the vertex set of A , and a new vertex q'_0 ;
- initial vertex q'_0 ;
- terminal vertices the terminal vertices of the automata D_{ij} ;
- all the edges of the automata C_{ij} and D_{ij} ;
- for each $i \in I'$ and $j \in J'$, an edge from q'_0 to the initial vertex of C_{ij} labelled $(1, \epsilon)$;
- for each $i \in I'$ and $j \in J'$, an edge from each terminal vertex of C_{ij} to (i, q_0, j) labelled $(1, \epsilon)$, where q_0 is the initial vertex for A ;
- for each edge in A from p to q labelled $((i, g, j), w)$ and each $i' \in I'$ and $j' \in J'$, an edge from (i', p, j') to (i', q, j) labelled $(P_{j'ig}, w)$;
- for each $i \in I', j \in J'$ and terminal vertex p of A , an edge from (i, p, j) to the initial vertex of D_{ij} labelled $(1, \epsilon)$.

Since I', J' and all the automata A, C_{ij} and D_{ij} are finite, it follows that the G -automaton B is finite. We now show that B accepts the language L .

Let $w \in L$. Then there exists a path through the automaton A labelled $((i, g, j), w)$ connecting the initial vertex with some terminal vertex (p_t say), such that

$$(i_0, g_0, j_0)(i, g, j) = (i', g', j') \in X_1$$

for some $(i_0, g_0, j_0) \in X_0$. Suppose this path has the form

$$q_0 \xrightarrow{((i_1, g_1, j_1), x_1)} q_1 \xrightarrow{((i_2, g_2, j_2), x_2)} q_2 \xrightarrow{((i_3, g_3, j_3), x_3)} \dots q_{m-1} \xrightarrow{((i_m, g_m, j_m), x_m)} q_m$$

where q_0 is the initial vertex and $q_m = p_t$ is a terminal vertex of A and $w = x_1 \dots x_m$. Note that we must have $i' = i_0, j' = j_m$ and

$$g = g_1 P_{j_1 i_2} g_2 \dots P_{j_{m-1} i_m} g_m.$$

Now by construction, B has a path π_2 of the form

$$(i_0, q_0, j_0) \xrightarrow{(P_{j_0 i_1} g_1, x_1)} (i_0, q_1, j_1) \xrightarrow{(P_{j_1 i_2} g_2, x_2)} (i_0, q_2, j_2) \xrightarrow{(P_{j_2 i_3} g_3, x_3)} \dots \xrightarrow{(P_{j_{m-1} i_m} g_m, x_m)} (i_0, q_m, j_m)$$

Moreover, from the fact that $(i_0, g_0, j_0) \in X_0$ we see that $g_0 \in X_{i_0 j_0}$, so that $(g_0, \epsilon) \in X'_{i_0 j_0}$. Hence, (g_0, ϵ) labels a path in $C_{i_0 j_0}$ from the initial vertex to a terminal vertex. It follows easily that (g_0, ϵ) labels a path π_1 in B from the initial vertex q'_0 to (i_0, q_0, j_0) where q_0 . Similarly, since $(i', g', j') \in X_1$ we deduce that $((g')^{-1}, \epsilon) \in Y'_{i' j'} = Y_{i_0 j_m}$ so that B has a path π_3 from (i_0, q_m, j_m) to a terminal vertex labelled $((g')^{-1}, \epsilon)$.

Composing the paths π_1, π_2 and π_3 , we see that B has a path from the initial vertex to a terminal vertex with label

$$(g_0 P_{j_0 i_1} g_1 P_{j_1 i_2} g_2 \dots P_{j_{m-1} i_m} g_m (g')^{-1}, x_1 x_2 \dots x_m).$$

But we know that $(i_0, g_0, j_0)(i, g, j) = (i', g', j')$, so we must have

$$g_0 P_{j_0 i_1} g_1 P_{j_1 i_2} g_2 \dots P_{j_{m-1} i_m} g_m = g'$$

and hence

$$g_0 P_{j_0 i_1} g_1 P_{j_1 i_2} g_2 \dots P_{j_{m-1} i_m} g_m (g')^{-1} = 1.$$

It follows that w is accepted by the G -automaton B , as required.

Conversely, suppose w is accepted by the G -automaton B . Then there is a path in B from the initial vertex to a terminal vertex labelled $(1, w)$. We deduce easily from the construction of B that this path must have the form $\pi_1 \pi_2 \pi_3$ where

- π_1 runs from the start vertex to some vertex (i_0, q_0, j_0) with label of the form (g_0, ϵ) for some $g_0 \in X_{i_0 j_0}$, so that $(i_0, g_0, j_0) \in X_0$;

- π_2 runs from (i_0, q_0, j_0) to a vertex (i_0, q_m, j_m) where q_m is a terminal vertex of A ; and
- π_3 runs from (i_0, q_m, j_m) to a terminal vertex with label $((g')^{-1}, \epsilon)$ where $(g')^{-1} \in Y_{i_0 j_m}$, so that $(i_0, g', j_m) \in X_1$.

Moreover, π_2 must have the form

$$(i_0, q_0, j_0) \xrightarrow{(P_{j_0 i_1} g_1, x_1)} (i_0, q_1, j_1) \xrightarrow{(P_{j_1 i_2} g_2, x_2)} (i_0, q_2, j_2) \xrightarrow{(P_{j_2 i_3} g_3, x_3)} \dots \xrightarrow{(P_{j_{m-1} i_m} g_m, x_m)} (i_0, q_m, j_m)$$

where, since the label of the entire path π is $(1, w)$, we must have $w = x_1 \dots x_m$ and $g_0 P_{j_0 i_1} g_1 \dots P_{j_{m-1} i_m} g_m (g')^{-1} = 1$, that is,

$$g_0 P_{j_0 i_1} g_1 \dots P_{j_{m-1} i_m} g_m = g'.$$

We deduce from the path above and the construction of B that A has a path

$$q_0 \xrightarrow{((i_1, g_1, j_1), x_1)} q_1 \xrightarrow{((i_2, g_2, j_2), x_2)} q_2 \xrightarrow{((i_3, g_3, j_3), x_3)} \dots q_{m-1} \xrightarrow{((i_m, g_m, j_m), x_m)} q_m$$

Since q_0 and q_m are initial and terminal vertices of A respectively, it follows that A accepts (x, w) where

$$x = (i_1, g_1, j_1)(i_2, g_2, j_2) \dots (i_m, g_m, j_m).$$

But (i_0, g_0, j_0) lies in X_0 and

$$\begin{aligned} (i_0, g_0, j_0)x &= (i_0, g_0, j_0)(i_1, g_1, j_1) \dots (i_m, g_m, j_m) \\ &= (i_0, g_0 P_{j_0 i_1} g_1 \dots P_{j_{m-1} i_m} g_m, j_m) \\ &= (i_0, g', j_m) \end{aligned}$$

lies in X_1 , from which we deduce that the S -automaton A accepts the word w , and so $w \in L$ as required. \square

Acknowledgement

The research of the second author was supported by an RCUK Academic Fellowship.

References

- [1] J. Berstel, *Transductions and Context-free Languages*, Teubner, 1979.
- [2] A.H. Clifford, G.B. Preston, *The Algebraic Theory of Semigroups*, vol. 1, American Mathematical Society, 1961.
- [3] G.Z. Elston, G. Ostheimer, On groups whose word problem is solved by a counter automaton, *Theoretical Computer Science* 320 (2–3) (2004).
- [4] H. Fernau, R. Stiebe, Valence grammars with target sets, in: S. Yu, M. Ito, Gh. Paun (Eds.), *Words, Semigroups and Transductions*, World Scientific, Singapore, 2001, pp. 129–140.
- [5] H. Fernau, R. Stiebe, Sequential grammars and automata with valences, *Theoretical Computer Science* 276 (2002) 377–405.
- [6] R.H. Gilman, Formal languages and infinite groups, in: *Geometric and Computational Perspectives on Infinite Groups* (Minneapolis, MN and New Brunswick, NJ, 1994), DIMACS Series, in: *Discrete Mathematics and Theoretical Computer Science*, vol. 25, American Mathematical Society, Providence RI, 1996.
- [7] R.H. Gilman, M. Shapiro, On groups whose word problem is solved by a nested stack automaton, 1998. [arXiv:math.GR/9812028](https://arxiv.org/abs/math/9812028).
- [8] J.M. Howie, *Fundamentals of Semigroup Theory*, Clarendon Press, 1995.
- [9] M. Kambites, Formal languages and groups as memory, *Communications in Algebra* 37 (2009) 193–208.
- [10] M. Kambites, Word problems recognisable by deterministic blind monoid automata, *Theoretical Computer Science* 362 (1) (2006) 232–237.
- [11] M. Kambites, P.V. Silva, B. Steinberg, On the rational subset problem for groups, *Journal of Algebra* 309 (2) (2007) 622–639.
- [12] V. Mitrana, R. Stiebe, The accepting power of finite automata over groups, in: *New Trends in Formal Languages*, in: *Lecture Notes in Computer Science*, vol. 1218, Springer, Berlin, 1997.
- [13] M. Nivat, Transductions des langages de Chomsky, *Université de Grenoble. Annales de l'Institut Fourier* 18 (1968) 339–455.
- [14] M. Nivat, Sur les automates à mémoire pile, in: W. Itzfeld (Ed.), *Proc. International Computing Symposium Bonn 1970*, North-Holland, 1970, pp. 655–663.
- [15] Gh. Paun, A new generative device: Valence grammars, *Revue Roumaine de Mathématiques Pures et Appliquées* XXV (6) (1980) 911–924.
- [16] D. Rees, On semi-groups, *Proceedings of the Cambridge Philosophical Society* 36 (1940) 387–400.
- [17] E. Render, M. Kambites, Polycyclic and bicyclic valence automata, in: *Proc. 2nd International Conference on Languages and Automata Theory and Applications*, LATA 2008, in: *LNCS*, vol. 5196, 2008, pp. 464–475.
- [18] E. Render, M. Kambites, Rational subsets of polycyclic monoids and valence automata, *Information and Computation* (in press).
- [19] J. Sakarovitch, Easy multiplications. I. the realm of Kleene's theorem, *Information and Computation* 74 (3) (1987) 173–197.