Preprint

DECIDING WORD PROBLEMS OF SEMIGROUPS USING FINITE STATE AUTOMATA

MAX NEUNHÖFFER, MARKUS PFEIFFER, AND NIK RUSKUC

ABSTRACT. We explore a natural class of semigroups that have word problem decidable by finite state automata. Among the main results are invariance of this property under change of generators, invariance under basic algebraic constructions and properties of the group of units.

1. MOTIVATION

Finite state automata are a simple concept that is well-established in the theory of computation. They are very restricted in that they only possess a finite memory. These restrictions cause many problems in the theory of finite state automata to be decidable and quite a few are tractable complexity-wise.

The *word problem* is a computational problem that is connected to finitely generated structures, especially finitely generated semigroups, monoids or groups. In this paper we want to explore the properties of semigroups that have word problem decidable by certain types of finite state automata.

There has been some research in this area by various authors, for example Kambites shows in [Kam09a] and [Kam09b] that semigroups with small cancellation have rational word problem, Holt et al. in [HOT08] investigate properties of semigroups with one-counter word problem.

Let us fix basic notation. Let A be a finite set. A string over A is a finite sequence of elements of A, and we denote the special case of the empty sequence by $''''_A$ or simply $''''$ if there is no ambiguity. We denote by A^* the set of all strings over A and by A^+ the set of all nonempty strings over A. Obviously $A^+ \subseteq A^*$. We denote by $|s|$ the length of a string s and by $|s|_a$ the number of occurrences of the letter a from A in s. In the following it will be important to distinguish between strings and elements of semigroups, monoids or groups, we write literal strings enclosed in quotation marks. Thus $''abc''$ is a string of length three in A^* for a, b, c in A in contrast to abc being the product of a, b and c in some structure containing A. The most important operation on strings is *concatenation*. Given two strings v and w, we denote the concatenation of v and w by $v.w$, which is just the juxtaposition of the two strings. Given a string v, we denote by v^i for any natural number i the $i-$fold concatenation of copies of v. The special case v^0 is defined to be $''''$.

A string s is a *prefix* of a string t if there is a string u in A^* such that $t = s.u$, and analogously s is a *suffix* of t if there is a string u such that $t = u.s$.

A *semigroup* S is a set together with a binary associative operation, which we usually denote by $s \cdot t$. We allow ourselves to leave out the dot in most cases. A *monoid* M is a semigroup which contains an identity element e for which $ea = ae = a$ for all a in M holds. A group G is a monoid with the additional condition that

---

Date: May 12, 2011.

for each g in G there is an element h such that $gh = hg = e$. We can make every semigroup into a monoid by adjoining an identity element. For a semigroup S we denote the semigroup with an adjoined identity element by S^1. Another important element in the theory of semigroups is a zero element, or zero for short. A zero z has the property that $za = az = z$ for all a in S. We can adjoin a zero to every semigroup and we denote this semigroup by S^0.

The set A^+ together with the concatenation operation is isomorphic to the free semigroup on A, which is uniquely defined by the following universal property: For a given semigroup S and any map p from A to S, the map p can be extended to a unique semigroup homomorphism from A^+ to S. Analogously the set A^* together with concatenation is isomorphic to the free monoid on A, which in turn is defined analogously.

In the following it is important to distinguish between strings over an alphabet representing elements of some semigroup, monoid or group and elements of the respective structure.

A semigroup S is finitely generated if there is a finite subset A of S such that the inclusion map from A to S extends to a *surjective* semigroup homomorphism $^- : A^+ \to S$. In this situation we write $S = \mathrm{Sg}\langle A \rangle$ to say that S is a semigroup that is finitely generated by the subset A of S. At this point it should be noted that although A is a subset of S, the set A^+ is a set of strings and is not a subset of S. If we want to pass over to the semigroup S we write $\overline{v}$ to denote the image of the string v in S.

Since monoids and groups are also semigroups the above defintion for semigroup generation works for groups and monoids too. In addition, we can also generate a monoid as a monoid by making the identity element implicit and we write $M = \mathrm{Mon}\langle A \rangle$ to say that the monoid M is generated as a monoid by the subset A of M. A group can be generated as a group by making inverses implicit.

We might for some examples use presentations to give semigroups or monoids and use $\mathrm{Sg}\langle A \mid R \rangle$ for the semigroup generated by A with relations R and $\mathrm{Mon}\langle A \mid R \rangle$ for the monoid generated by A with relations R. For a reference on presentations the reader is referred to standard literature, for example [CP67].

For a finitely generated semigroup S and a generating set A for S we define the *semigroup word problem* to be the following set

$$\mathrm{SgWP}(S, A) := \left\{(v, w) \in A^+ \times A^+ \mid \overline{v} = \overline{w}\right\} \subseteq A^+ \times A^+.$$

For a finitely generated monoid M and some generating set A for M we define the *monoid word problem* to be the following set

$$\mathrm{MonWP}(M, A) := \{(v, w) \in A^* \times A^* \mid \overline{v} = \overline{w}\} \subseteq A^* \times A^*.$$

This is in close relation to the usual definition of the word problem for a finitely generated group G with respect to a finite semigroup generating set A

$$\mathrm{GrpWP}(G, A) := \left\{v \in (A \cup A^{-1})^* \mid \overline{v} = e\right\} \subseteq (A \cup A^{-1})^*,$$

because if G is a finitely generated group then $\mathrm{SgWP}(G, A)$ consists of pairs (v, w) such that $\overline{v} \cdot \overline{w}^{-1} = e$.

Note that we defined three word problems for any given finitely generated group G, and two word problems for any given finitely generated monoid M.

A fundamental computational question that arises in this context is whether the word problems of finitely generated semigroups, monoids and groups are *decidable* subsets of $A^* \times A^*$ or $(A \cup A^{-1})^*$ respectively. That is, do there exist algorithms that take as input elements from $A^* \times A^*$ or $(A \cup A^{-1})^*$ respectively and that terminate with the output **true** or **false** depending on whether the input is contained in the respective word problem or not. A finite state automaton can be seen as a very simple algorithm, in particular one that does requires constant memory.

In the following we want to explore the properties of semigroups, monoids and groups with word problem decidable by a finite state automaton. The results for groups are already well-known, results concerning semigroups and monoids are original work of the authors.

The paper is structured as follows: Section 2 will introduce the basic theory we want to use in proving our results. In particular, we define what we mean by a semigroup with regular or rational word problem. After that, Section 3 presents basic and motivational results and in Section 4 we prove the first main result, invariance of rational word problem under change of generating sets, and Section 5 establishes the related result that a finitely generated monoid has rational word problem generated as a semigroup if and only if it has rational word problem generated as a monoid. Section 6 establishes a few structural properties of semigroups with rational word problem, and Sections 7 and 8 then deal with constructions involving semigroups with rational word problem and closure properties. Section 9 will give a few pointers towards followup research and papers.

2. AUTOMATA

{sec:prelim}

In this section we recall the definitions of finite state automata that take strings as input, and two tape finite state automata which take pairs of strings as input. In addition to the basic definitions we give some results from the theory of finite state automata for later reference.

Definition 2.1 (finite state automaton)**.** *A* finite state automaton $\mathfrak{A}$ *is a tuple*

$$\mathfrak{A} = \langle Q, A, q_0, F, \Delta \rangle$$

consisting of a finite set Q of states, an alphabet A, an initial state q_0 in Q, a set $F \subseteq Q$ of final states and a transition relation $\Delta \subseteq Q \times (A \cup \{\varepsilon\}) \times Q$.

We also denote elements (q, a, r) from Δ by

$$q \xrightarrow{a} r.$$

A *computation* of $\mathfrak{A}$ from q_1 to q_{n+1} with *label* "$a_1 a_2 \cdots a_n$" is a finite sequence of transitions

$$\gamma : q_1 \xrightarrow{a_1} q_2 \xrightarrow{a_2} q_3 \xrightarrow{a_3} \cdots \xrightarrow{a_{n-1}} q_n \xrightarrow{a_n} q_{n+1}.$$

The computation γ is said to be *accepting* if q_1 is the initial state and q_{n+1} is an element of F. Note that the label of a computation is an element of $(A \cup \{\varepsilon\})^*$.

Consider the map

$$p : (A \cup \{\varepsilon\}) \to A^* : a \mapsto \begin{cases} "a" & \text{for } a \in A \\ "" & \text{for } a = \varepsilon \end{cases},$$

which extends to a surjective monoid homomorphism $\pi : (A \cup \{\varepsilon\})^* \to A^*$.

We say that $\mathfrak{A}$ *accepts* a string s in A^* if there is an accepting computation labelled by a string t in $(A \cup \{\varepsilon\})^*$ such that $\pi(t) = s$. The set of all strings in A^* that are accepted by $\mathfrak{A}$ is called the *language* of $\mathfrak{A}$, denoted $L(\mathfrak{A})$.

Conversely, subsets L of A^* with $L = L(\mathfrak{A})$ for some finite state automaton $\mathfrak{A}$ are called *regular*.

A slight generalisation of the concept of a finite state automaton is the notion of a *synchronous two tape finite state automaton*. For this we take an alphabet A and add a padding character $\square$ forming $A^{\square} := A \cup \{\square\}$. As alphabet for a two tape synchronous finite state automaton we take $A^{\square} \times A^{\square}$. To be able to feed pairs (s,t) from $A^* \times A^*$ of strings of differing length to such an automaton we pad the shorter of the two strings by using the padding symbol, more formally

$$(s,t)^{\square} := "(s'_1, t'_1)(s'_2, t'_2) \cdots (s'_n, t'_n)",$$

where $n = \max\{|s|, |t|\}$ and

$$z'_i = \begin{cases} z_i & i \leq |z| \\ \square & \text{otherwise} \end{cases}$$

for $1 \leq i \leq n$ and $z \in \{s,t\}$.

We call a subset R of $A^* \times A^*$ *regular* if there is a synchronous two tape finite state automaton that accepts a padded pair $(s,t)^{\square}$ if and only if (s,t) is in R.

Note that $(A \times B)^*$ is isomorphic to the submonoid of pairs of strings of equal lengths in $A^* \times B^*$ and we will use this isomorphism implicitly.

Generalising further, an *asynchronous two tape finite state automaton* has the ability to read its two tapes at different speeds.

Definition 2.2 (asynchronous finite state automaton)**.** *An* asynchronous finite state automaton $\mathfrak{A}$ *is a tuple*

$$\mathfrak{A} := \langle Q, A, B, q_0, F, \Delta \rangle$$

consisting of a finite set Q of states, two alphabets A and B, an initial state q_0 in Q, a set $F \subseteq Q$ of final states and a transition relation $\Delta \subseteq Q \times (A \cup \{\varepsilon\}) \times (B \cup \{\varepsilon\}) \times Q$

Analogous to the case of a finite state automaton, we denote elements (p,a,b,q) of the transition relation by

$$p \xrightarrow{(a,b)} q,$$

and a computation γ of $\mathfrak{A}$ from q_1 in Q to q_{n+1} in Q with label $"(a_1,b_1) \cdots (a_n,b_n)"$ is a finite sequence of transitions, denoted

$$\gamma : q_1 \xrightarrow{(a_1,b_1)} q_2 \xrightarrow{(a_2,b_2)} q_3 \cdots q_n \xrightarrow{(a_n,b_n)} q_{n+1}.$$

We shorten this to $\gamma : q_1 \to^* q_{n+1}$ to say that there is a computation of finite length from q_1 to q_{n+1}. A computation γ is said to be accepting if $q_1 = q_0$ and q_{n+1} is in F.

In the case of an asynchronous automaton the label of a computation is an element of $((A \cup \{\varepsilon\}) \times (B \cup \{\varepsilon\}))^*$.

To get a pair of strings from the label of a computation we apply maps π_A and π_B analogous to the case of finite state automata to both components of the pair of strings that arises from the label of the computation. We also say that a pair (v,w) of strings *induces* a computation $\gamma : q_1 \to^* q_n$ if γ has label (s,t) such that $(\pi_A(s), \pi_B(t)) = (v,w)$.

An asynchronous automaton $\mathfrak{A}$ is said to accept a pair (s,t) of strings in $A^* \times B^*$ if there is an accepting computation of $\mathfrak{A}$ with label (v,w) such that $(\pi_A(v), \pi_B(w)) = (s,t)$.

The set of all pairs (s,t) that are accepted by a finite state automaton $\mathfrak{A}$ is called the language of $\mathfrak{A}$ and is denoted $L(\mathfrak{A})$.

Subsets R of $A^* \times B^*$ for which there is an asynchronous finite state automaton $\mathfrak{A}$ with $L(\mathfrak{A}) = R$ are called *rational relations* or simply *rational*.

For any of the above automaton models, we call a state q in $\mathfrak{A}$ *accessible* if there is a computation in $\mathfrak{A}$ from the initial state q_0 to q and *co-accessible* if there is computation from q to a final state. An automaton is *unambigious* if for any string s and any pair p and q of states there is at most one computation from p to q induced by s. Furthermore an automaton $\mathfrak{A}$ is *deterministic*, if it is unambigous and for any given input there is at least one computation that is induced by that input.

The above automaton models have a natural interpretation as finite, directed, labelled graphs where the set of vertices is the set of states and there is a labelled edge between two states if and only if there is a transition between them.

We now recall the well known Pumping Lemmas which enable us to prove that a set is not regular or rational respectively. For proofs of the two lemmas we refer the reader to [HMU01] for the finite state automaton case and to [Ber79] for the asynchronous finite state automaton case. In fact, the proof of the asynchronous case uses the synchronous one.

{prop:pump}

Proposition 2.3 (Pumping Lemma for finite state automata)**.** *Let $\mathfrak{A}$ be a finite state automaton. Then there is a natural number n_0 such that for every string s accepted by $\mathfrak{A}$ with $|s| > n_0$ there is a decomposition $s = x.u.y$ into strings x, u and y such that*

- $|u| \geq 1$
- $|x.u| \leq n_0$
- *For all $i \in \mathbb{N}$ the string $x.u^i.y$ is also accepted by $\mathfrak{A}$.*

Note that there are languages that are not regular but fulfill the Pumping Lemma. In our context the Pumping Lemma is useful to show that a language cannot be accepted by a finite state automaton.

{prop:async}

Proposition 2.4 (Pumping Lemma for asynchronous finite state automata)**.** *Let $\mathfrak{A}$ be an asynchronous finite state automaton. Then there is a natural number n_0 such that for every pair (s_1, s_2) of strings accepted by $\mathfrak{A}$ with $|s_1| + |s_2| > n_0$ there is a decomposition $(s_1, s_2) = (x_1.u_1.y_1, x_2.u_2.y_2)$ into pairs (x_1, x_2), (u_1, u_2) and (y_1, y_2) such that*

- $|u_1| + |u_2| \geq 1$
- $|x_1| + |x_2| + |u_1| + |u_2| \leq n_0$
- *For all $i \in \mathbb{N}$ the pair $(x_1.u_1^i.y_1, x_2.u_2^i.y_2)$ is also accepted by $\mathfrak{A}$.*

The following proposition states that the composition of rational relations is again a rational relation.

{prop:ratcomp}

Proposition 2.5. *Let A,B and C be alphabets and let $R \subseteq A^* \times B^*$ and $S \subseteq B^* \times C^*$ be rational relations. Then $R \circ S$ is also a rational relation, where*

$$R \circ S = \{(r, s) \in A^* \times C^* \mid \text{there is } x \in B^* \text{ such that } (r, x) \in R \text{ and } (x, s) \in S\ \}$$

Proof. See [Ber79]. □

J.H. Johnson in his PhD thesis [Joh83] examined rational equivalence relations over strings, that is rational relations that are equivalence relations. He proved the following theorem which we will use in a later section to show that infinite semigroups with rational word problem cannot be periodic. The proof can be found in the referenced paper.

{prop:rateqthin}

Proposition 2.6. *Let A be an alphabet and $R \subseteq A^* \times A^*$ be a rational equivalence relation. Then there is a regular language $D \subseteq A^*$ that contains at least one element of each equivalence class of R and is such that $R \cap (D \times D)$ is a rational equivalence relation on D*

Proof. The idea of the proof is to remove loops from an automaton that decides R that are labelled by (s, ε) for some $s \in (A \cup \{\varepsilon\})^*$ to make it accept long representatives. The language of long representatives is regular and thus its complement is language D. See [Joh85]. □

The following two propositions will help simplify the proofs of a few theorems. The proofs are straightforward and can be found in [Ber79].

:directproduct_reg_rat}

Proposition 2.7. *Let A and B be two alphabets. If L_1 is a regular language over A and L_2 is a regular language over B, then $L_1 \times L_2$ is a rational relation.*

{prop:preimage_regular}

Proposition 2.8. *Let A and B be alphabets and $R \subseteq A^* \times B^*$ be a rational relation. Then the languages $\{w \in B^* \mid (v, w) \in R\}$ and $\{v \in A^* \mid (v, w) \in R\}$ are regular for all $v \in A^*$ and $w \in B^*$.*

{sec:firstres}

3. MOTIVATING RESULTS AND EXAMPLES

This section is dedicated to motivate the presented work by starting from a well known result about groups: Anisimov showed in [Ani71] that the class of groups with regular word problem is the class of finite groups. We show that finite semigroups and monoids have regular word-problem as well and give a very simple example of a semigroup with regular word-problem that is infinite. We also motivate the usage of rational word-problem for infinite semigroups and monoids. In Section 5 we show that groups with rational word-problem are finite and thus our theory does not yield more expressiveness for groups.

{gp_regwp_iff_finite}

Theorem 3.1 (Anisimov)**.** *Let G be a group and let A be a finite monoid generating set for G. Then* $\mathrm{GrpWP}(G, A) \subseteq A^*$ *is regular if and only if G is finite.*

Proof. Suppose G is finite and consider the automaton

$$\mathfrak{A} = \langle G, A, 1, \{1\}, \Delta \rangle,$$

where (g, a, h) is in Δ if and only if $ga = h$. This automaton is the Cayley graph of G with respect to the generating set A extended by predicates for the initial state and final states.

A string s in A^* is accepted by $\mathfrak{A}$ if and only if there is a computation from 1 to 1 labelled by s. This also means that $\overline{s} = 1$ by the definition of $\mathfrak{A}$.

Conversely, assume that there is a finite state automaton $\mathfrak{A} = \langle Q, A, q_0, F, \Delta \rangle$ that has as its language all strings s with $\overline{s} = 1$. Without loss of generality we can assume $\mathfrak{A}$ to be deterministic, because if it was not we can construct an equivalent deterministic automaton by applying the powerset construction, which is a standard tool in the theory of finite state automata and can for example be found in [HMU01].

Let s and t be two strings that label paths in $\mathfrak{A}$ from q_0 to some state q in Q. Since G is a group and $\mathfrak{A}$ is deterministic there has to be a path from q labelled u to an accept state. Thus

$$\overline{s.u} = \overline{s} \cdot \overline{u} = 1 = \overline{t} \cdot \overline{u} = \overline{t.u},$$

which implies $\overline{s} = \overline{t}$ and therefore G is finite. □

One direction of the above theorem stays true for semigroup and monoid word-problems.

{thm:finmon_regular_wp}

Theorem 3.2. *Let S be a finite semigroup or monoid. Then S has regular word-problem with respect to all generating sets.*

Proof. Let S be a finite semigroup and let A be any generating set for S. Consider the following automaton.

$$\mathfrak{A} = \left\langle Q, A^{\square} \times A^{\square}, q_0, F, \Delta \right\rangle$$

consisting of

$$\begin{array}{rcl} Q & = & \{q_0\} \cup (S \times S \times \{L, N, R\}) \\ F & = & \{(s, s, N) \mid s \in S\} \\ \Delta & = & \{[q_0, (x, y), (x, y, N)]\} \\ & & \cup\ \{[(s, t, i), (x, y), (sx, ty, i)] \mid x, y \neq \square, i \in \{L, N, R\}\} \\ & & \cup\ \{[(s, t, N), (x, \square), (sx, t, R)]\} \\ & & \cup\ \{[(s, t, N), (\square, y), (s, ty, L)]\} \\ & & \cup\ \{[(s, t, R), (x, \square), (sx, t, R)]\} \\ & & \cup\ \{[(s, t, L), (\square, y), (s, ty, L)]\}\,. \end{array}$$

This automaton consists of three copies of the direct product of two copies of the Cayley graph of S together with an initial state. Reading a pair of symbols it keeps track of right multiplication by a generator with the $\square$ symbol acting as identity. This way, the automaton determines the elements represented by the input strings and accepts if and only if these are the same.

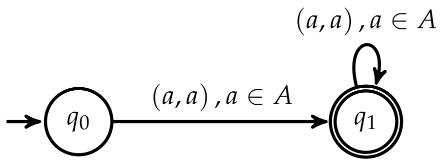
The copies indexed by L, N and R are needed to take care of padding symbols: if a padding symbol is read on one tape for the first time the automaton is only allowed to read padding symbols from that tape and non padding symbols from the other tape.

This automaton accepts a pair $(v, w)^{\square}$ of padded strings if and only if $\overline{v} = \overline{w}$. The proof for monoids is similar. $\square$

In contrast with the group case there are examples of infinite semigroups and monoids that do have regular word problem. The most striking examples are the free semigroup and the free monoid on any finite set. Additionally, in Sections 7 and 8 it will be shown that we can construct infinite semigroups with rational word problem from semigroups which are known to have rational word problem.

Example 3.3. *Let A be a finite, non-empty set. Then the free semigroup A^+ and the free monoid A^* are infinite and* $\mathrm{SgWP}(A^+, A)$ *and* $\mathrm{MonWP}(A^*, A)$ *are regular.*

Proof. The following automaton accepts pairs of equal strings.



In a free semigroup on a finite set A two strings v and w represent the same element if and only if they are equal, therefore $\mathrm{SgWP}(A^+, A)$ is regular. For an automaton that decides $\mathrm{MonWP}(A^*, A)$ we turn q_0 into an accept state. $\square$

One important aspect in the above definitions of the word problem is the dependence on the generating set. In general, if for a semigroup S with generating set A the set $\mathrm{SgWP}(S, A)$ is regular, then this might not be true for other finite generating sets of S. The following theorem characterises the semigroups that have regular word problem with respect to every generating set.

{thm:regular_iff_finite

Theorem 3.4. *Let S be a finitely generated semigroup. Then* $\mathrm{SgWP}(S, A)$ *is regular for all finite generating sets A if and only if S is finite.*

Proof. The if part is precisely Theorem 3.2.

Suppose S is infinite. Then by Theorem 6.1 there exists some s in S that has infinite order. Let A be a generating set for S. The set $B := A \cup \{s, t\}$ where $t = s^2$ also generates S. Applying the Pumping Lemma to the pair $(t^{n_0}, s^{2n_0})^{\square}$ shows that the set $\mathrm{SgWP}(S, B)$ is not regular. ☐

A consequence of the preceding paragraph is that $\mathrm{SgWP}(S, A)$ being regular depends on the choice of the generating set for infinite semigroups. We will show in Section 4 that using asynchronous finite state automata is the appropriate choice of automaton model to achieve independence of change of generators while keeping a finite state device.

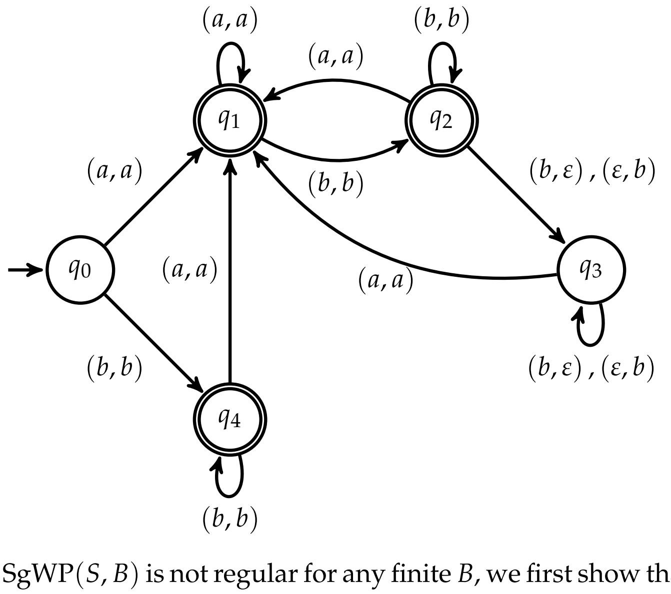
The following example shows that there are semigroups that are finitely generated, not finitely presentable and have rational word problem.

am:non_fin_pres_rat_wp}

Example 3.5. *Let* $S = \mathrm{Sg}\,\langle a, b \mid (ab^n a = aba)_{n \geq 2}\rangle$. *This semigroup is infinite, not finitely presentable and* $\mathrm{SgWP}(S, \{a, b\})$ *is rational. Furthermore there is no generating set A' for S such that* $\mathrm{SgWP}(S, A')$ *is regular.*

Proof. The monoid S is infinite because the submonoid generated by a is infinite. If S had a finite presentation then there would be a set $X \subseteq \{ab^n a = aba \mid n \geq 2\}$ such that $S \cong \mathrm{Sg}\,\langle a, b \mid X\rangle$. This would mean that there is an $N \in \mathbb{N}$ such that $ab^N a = aba$ is a consequence of $ab^k a = aba$ for k less than N, which is impossible.

To show that S has rational word problem we give the following asynchronous finite state automaton that decides the word problem of S.



To show that $\mathrm{SgWP}(S, B)$ is not regular for any finite B, we first show that $\mathrm{SgWP}(S, A)$ is not regular. For this assume $\mathrm{SgWP}(S, A)$ to be regular and to be accepted by a finite state automaton with n states. Choose $n_0 > n$ and consider the pair

$$\left(''ab''^{n_0}.''a''.''b''^{2n_0}.''a'', ''a''.''b''^{2n_0}.''ab''^{n_0}.''a'' \right).$$

Both strings represent the same element $(ab)^{n_0+1}a$ of S and therefore the pair is an element of $\mathrm{SgWP}(S, A)$.

Since $n_0 > n$ there are two natural numbers i and j with $i < j$ such that after reading $(''ab''^i, ''a''.''b''^{2i-1})$ and $(''ab''^j, ''a''.''b''^{2j-1})$ the automaton is in some state q. From $(''ab''^i, ''a''.''b''^{i-1})$ the automaton can reach an accept state by reading the pair $(''a''.''b''^{2i-3}.''a'', ''ab''^{i-1}.''a'')$. Hence the automaton also accepts

$$(''ab''^j.''a''.''b''^{2i-3}.''a'', ''a''.''b''^{2j-1}.''ab''^{i-1}.''a'')$$

which would mean that $(ab)^{j+1}a$ is equal to $(ab)^{i+1}a$ in contradiction to $j > i$.

Any generating set for S must give rise to representatives for a and b and thus this argument also holds for any generating set of S. Therefore there cannot be a generating set B for S such that $\mathrm{SgWP}(S, B)$ is regular. □

The following lemma shows that a free commutative semigroup of rank at least two does not have rational word problem. This will also become one way of showing that a semigroup does *not* have rational word problem.

{exam:nonregular}

Lemma 3.6. *Let* $A = \{a, b\}$ *and* $S = \mathrm{Sg}\langle A \mid ab = ba\rangle$. *Then* $\mathrm{SgWP}(S, A)$ *is not regular.*

Proof. For completeness we demonstrate how the pumping lemma is useful in this context.

Two strings s and t over A represent the same element of M if and only if $|s|_a = |t|_a$ and $|s|_b = |t|_b$.

For a contradiction assume that $\mathrm{SgWP}(M, A)$ is regular. By the Pumping Lemma there exists a natural number n_0 such that for all strings $s \in \mathrm{SgWP}(S, A)$ with $|s| > n_0$ there is a factorisation $s = x.u.y$ of s with $|x.u| < n_0$ such that $x.u^i.y$ is also in $\mathrm{SgWP}(S, A)$ for all $i \in \mathbb{N}$.

Consider the two representatives $''a''^{n_0}.''b''^{n_0}$ and $''b''^{n_0}.''a''^{n_0}$ of the same element of S. Thus $s = (''a''^{n_0}.''b''^{n_0}, ''b''^{n_0}.''a''^{n_0})$ is an element of $\mathrm{SgWP}(S, A)$. Since $|s| = 2n_0 > n_0$, there is a factorisation $s = x.u.y$ of s with $|x.u| \leq n_0$ such that $x.u^i.y$ is also in $\mathrm{SgWP}(S, A)$. The factors are

- $x = (''a''^k, ''b''^k)$
- $u = (''a''^l, ''b''^l)$
- $y = (''a''^{n_0-k-l}.''b''^{n_0}, ''b''^{n_0-k-l}.''a''^{n_0})$ for $k, l \in \mathbb{N}$ with $k + l < n_0$.

But $''a''^k.''a''^{il}.''b''^{n_0-k-l}$ and $''b''^k.''b''^{il}.''a''^{n_0-k-l}$ do not represent equal elements in S for $i > 1$, thus $\mathrm{SgWP}(S, A)$ is not regular, not even rational by Proposition 2.3. □

As a closing example for this section, we show that a very important and well known monoid does not have rational word problem.

{lem:bicyc_not_rat}

Lemma 3.7. *The bicyclic monoid* $B = \mathrm{Mon}\langle b, c \mid bc = 1\rangle$ *does not have rational word problem.*

Proof. This can be proven by applying the Pumping Lemma to $(''b''^{n_0}''c''^{n_0}, '''')$ for an appropriate n_0. □

4. CHANGE OF GENERATORS AND SUBSEMIGROUPS

{sec:change_of_generato

This section is dedicated to showing that for finitely generated semigroups rational word problem is independent of the choice of a finite generating set. The proof employs closure of rational relations under composition.

To prove the main result of this section we first give a few technical lemmas. We observe that the graph of a map that replaces every occurrence of some symbol in a string by a string is a rational relation, after that we show how closure of rational relations under compositions helps proving the main theorem.

{lem:ratmap}

Lemma 4.1. *Let A be an alphabet and $B = A \cup \{b\}$, where b is not an element of A. For some string w over A consider the following map:*

$$\varphi : B \to A^*, x \mapsto \begin{cases} w & \text{if } x = b \\ x & \text{otherwise} \end{cases} .$$

This map extends to a surjective morphism $\Phi : B^* \to A^*$ that replaces all occurrences of b in a string over B with w. The sets

$$R := \{(v, \Phi(v)) \in B^* \times A^* \mid v \in B^*\}$$

and

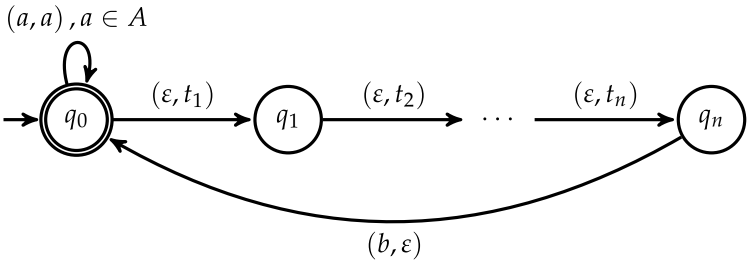
$$R^r := \{(\Phi(v), v) \in A^* \times B^* \mid v \in B^*\}$$

are rational relations.

Proof. Let $w = "t_1 \dots t_n"$ and consider $\mathfrak{R} = \langle Q, B, A, q_0, F, \Delta \rangle$, where

$$\begin{array}{rcl} Q & = & \{q_0, \dots, q_n\} \\ F & = & \{q_0\} \\ \Delta & = & \{(q_0, a, a, q_0) \mid a \in A\} \\ & & \cup\ \{(q_{i-1}, \varepsilon, t_i, q_i) \mid 1 \le i \le n\} \\ & & \cup\ \{(q_n, b, \varepsilon, q_0)\} \end{array}$$

Note that the states of $\mathfrak{R}$ correspond to prefixes of w. A picture makes the situation much easier to understand.



This automaton decides R. ☐

This next lemma shows how composition of rational relations helps us.

{la:equal}

Lemma 4.2. *Let S be a semigroup generated by the finite set A and let $B = A \cup \{b\}$, where b is an element of S not in A. Choosing w in A^+ such that $\overline{w} = b$, define R and R^r as in Lemma 4.1. Then the word problem* $\mathrm{SgWP}(S, B)$ *can be written in terms of R, R^r and* $\mathrm{SgWP}(S, A)$ *as follows:*

$$\mathrm{SgWP}(S, B) = R \circ \mathrm{SgWP}(S, A) \circ R^r.$$

If $\mathrm{SgWP}(S, A)$ *is rational then so is* $\mathrm{SgWP}(S, B)$.

Proof. Note that for all $u \in B^*$ the equality $\overline{u} = \overline{\Phi(u)}$ holds. Therefore for all $(v, w) \in A^+ \times A^+$

$$(v, w) \in \mathrm{SgWP}(S, B) \Leftrightarrow (\Phi(v), \Phi(w)) \in \mathrm{SgWP}(S, A)\,.$$

Also observe that for all $(u, w) \in B^+ \times A^+$

$$(u, w) \in R \Leftrightarrow w = \Phi(u) \Leftrightarrow (w, u) \in R^r.$$

Therefore

$$\begin{aligned} & (v,w) \in \mathrm{SgWP}(S,B) \\ \Leftrightarrow\quad & (v,\Phi(v)) \in R, (\Phi(w),w) \in R^r \text{ and } (\Phi(v),\Phi(w)) \in \mathrm{SgWP}(S,A) \\ \Leftrightarrow\quad & \exists v',w' \in A^* \text{ such that } (v,v') \in R, (w',w) \in R^r \text{ and } (v',w') \in \mathrm{SgWP}(S,A) \\ \Leftrightarrow\quad & (v,w) \in R \circ \mathrm{SgWP}(S,A) \circ R^r. \end{aligned}$$

It follows from Proposition 2.5 and Lemma 4.1 that if $\mathrm{SgWP}(S,A)$ is rational, then $\mathrm{SgWP}(S,B)$ is rational as well. ☐

The preceding lemmas are tied together to form the following theorem.

{thm:change_of_generato

Theorem 4.3. *Let S be a semigroup finitely generated by A such that $\mathrm{SgWP}(S,A)$ is rational.*

{thm:change_of_generato

{thm:change_of_generato

1. *If $B := A \cup \{b\}$ where b is an element of S not in A, then $\mathrm{SgWP}(S,B)$ is rational.*
2. *For the subsemigroup S' generated by $C := A\setminus\{c\}$ for any $c \in A$ the word problem $\mathrm{SgWP}(S',C)$ is rational.*

Proof. To prove (1), the relation $\mathrm{SgWP}(S,B)$ can be decomposed as shown in Lemma 4.2 and is rational. For (2) assume $\mathfrak{A} = \langle Q, A, A, q_0, F, \Delta\rangle$ to be the asynchronous finite state automaton that decides $\mathrm{SgWP}(S,A)$, then by removing all transitions involving a results in a new automaton that decides $\mathrm{SgWP}(S',C)$. ☐

The promised results for this section are now corollaries of Theorem 4.3.

{cor:fg_change_of_gens}

Corollary 4.4. *Let S be a semigroup. If there exists a finite generating set A for S such that $\mathrm{SgWP}(S,A)$ is rational then for all finite generating sets B of S the set $\mathrm{SgWP}(S,B)$ is rational.*

Proof. If there is a generating set A such that $\mathrm{SgWP}(S,A)$ is rational and given any other generating set B for S, we add generators from $B\setminus A$ to A using Theorem 4.3 (1). Then we remove everything in $A\setminus B$ by application of Theorem 4.3 (2). ☐

{cor:fg_subsemigroups}

Corollary 4.5. *Let S be a semigroup finitely generated by A and with $\mathrm{SgWP}(S,A)$ rational. Then for every finitely generated subsemigroup T the word problem $\mathrm{SgWP}(T,A')$ is rational.*

Proof. Given any finitely generated subsemigroup T of S, this follows from 4.3 by first adding a generating set for T to A and then removing superflous generators from the resulting set. ☐

The preceding corollaries give us a means of proving non-rationality of the word problem of semigroups. For example, if a semigroup contains a free commutative semigroup of rank greater than one, it cannot have rational word problem. This leads into Section 6 in which we discuss structural properties of semigroups with rational word problem.

5. CHANGE OF TYPE

{sec:change_of_type}

In this section we will show that groups with rational word problem are finite, and that monoids have rational word problem regardless of whether they are generated as a semigroup or a monoid.

We show that using asynchronous finite state automata does not result in more power for groups: A group G with rational monoid word problem is still finite. We will extend this result further by showing that the group of units of a monoid is finite in Theorem 6.2 and by showing that in fact any group contained in a semigroup with rational word problem has to be finite in Theorem 6.3.

{thm:group_as_monoid_fi

Theorem 5.1. *Let G be group, generated by a finite set A as a monoid. Then the monoid word problem* $\mathrm{MonWP}(G, A)$ *is rational if and only if G is finite.*

Proof. If G is finite, it follows from Theorem 3.2 that the $\mathrm{MonWP}(G, A)$ is regular and thus rational.

Suppose that G is an infinite group finitely generated by A and that $\mathfrak{A} = \langle Q, A, A, q_0, F, \Delta\rangle$ an asynchronous finite state automaton that decides $\mathrm{MonWP}(G, A)$. Without loss of generality assume $\mathfrak{A}$ to be accessible and co-accessible because any state not reachable from the initial state and every state from which no final state can be reached cannot occur in an accepting computation and can be removed without changing the accepted relation.

Let (v_1, w_1) and (v_2, w_2) be two pairs of strings that induce computations $\gamma_1 : q_0 \to^* q$ and $\gamma_2 : q_0 \to^* q$ respectively for some fixed state $q \in Q$. The quotients $\overline{w_1}^{-1}\overline{v_1}$ and $\overline{w_2}^{-1}\overline{v_2}$ coincide, because q is co-accessible and there is a pair (s, t) that induces a computation $\delta : q \to^* q_f$ to some accept state q_f, because G is a group. Therefore $\overline{v_1.s} = \overline{w_1.t}$ and $\overline{v_2.s} = \overline{w_2.t}$ which after rearrangement yields

$$\overline{w_1}^{-1}\overline{v_1} = \overline{t}\overline{s}^{-1} = \overline{w_2}^{-1}\overline{v_2}.$$

In particular if there are computations $\gamma_1 : q_0 \to^* q$ and $\gamma_2 : q_0 \to^* q$ induced by $(v_1, "\varepsilon"^{|v_1|})$ and $(v_2, "\varepsilon"^{|v_2|})$ respectively then $\overline{v_1} = \overline{v_2}$. Since G is infinite, there have to be two strings w_1 and w_2 such that $\overline{w_1} \neq \overline{w_2}$ and such that $(w_1, "\varepsilon"^{|w_1|})$ and $(w_2, "\varepsilon"^{|w_2|})$ induce computations $\gamma_1 : q_0 \to^* q$ and $\gamma_2 : q_0 \to^* q$ for some state q. This contradicts the choice of w_1 and w_2. □

Moving from semigroup generation to monoid generation and vice versa is possible for monoids without destroying rational word problem.

Theorem 5.2. *Let M be a finitely generated monoid and let A be a finite monoid generating set for M. Let* $S = \mathrm{Sg}\langle A\rangle$ *be the subsemigroup of M generated by A. Then the semigroup word-problem* $\mathrm{SgWP}(S, A)$ *is rational if and only if the monoid word-problem* $\mathrm{MonWP}(M, A)$ *is rational.*

Proof. Let M be a finitely generated monoid and let A be a finite monoid generating set for M that does not contain the identity element of M and let $S = \mathrm{Sg}\langle A\rangle$. Suppose that $\mathrm{SgWP}(S, A)$ is rational. The set

$$E = \left\{v \in A^+ \mid \overline{v} = e\right\},$$

where e is the identity element of M is regular, because if $e \in \mathrm{Sg}\langle A\rangle$ then there is a string w over A with $\overline{w} = e$ and thus E is regular by Proposition 2.8, and if $e \notin \mathrm{Sg}\langle A\rangle$ then E is empty. Therefore the set

$$W = \mathrm{SgWP}(S, A) \cup (E \times \{\varepsilon\}) \cup (\{\varepsilon\} \times E) \cup \{(\varepsilon, \varepsilon)\}$$

is rational and in fact $W = \mathrm{MonWP}(M, A)$.

Conversely, assume $\mathrm{MonWP}(M, A)$ is rational. We observe that

$$\mathrm{SgWP}(S, A) = \mathrm{MonWP}(M, A) \cap \left(A^+ \times A^+\right).$$

It remains to be shown that the intersection on the left hand side is rational. For this we use that the intersection of a rational and a recogniseable subset of a monoid is rational. This result can be found in [Ber79]. Since $\mathrm{MonWP}(M, A \cup \{e\})$ is rational and $A^+ \times A^+$ as a subset of $A^* \times A^*$ is recogniseable, the result follows. □

:structural_properties}

6. STRUCTURAL PROPERTIES

Having proven in Section 4 that $\mathrm{SgWP}(S, A)$ being rational is independent of the choice of A and thus a property of S, we want to establish structural results about such semigroups. We prove that semigroups with rational word problem cannot be periodic, monoids with rational word problem have finite group of units and that in fact all groups contained in a semigroup with rational word problem have to be finite.

We first show that an infinite semigroup with rational word problem cannot be periodic.

{thm:rat_not_periodic}

Theorem 6.1. *Let S be an infinite semigroup with rational word problem. Then there is an element y such that the subsemigroup $\mathrm{Sg}\,\langle y\rangle$ of S is infinite.*

Proof. Proposition 2.6 ensures existence of a regular language D that contains only finitely many representatives for each element of S. Since D is regular, Proposition 2.3 implies the existence of a natural number n_0 such that for every v in D with $|v| > n_0$ there exists a factorisation of v into three substrings x, y and z, such that $|y| \geq 1$, $|x.y| < n_0$ and $x.y^i.z \in D$ for all $i \in \mathbb{N}$. This means that the element $\overline{y}$ has to have infinite order. ☐

For monoids an interesting submonoid is the group of units. A monoid with rational word problem can only contain a finite group of units as is shown in the following theorem.

{thm:group_of_unity_fin

Theorem 6.2. *Let M be a finitely generated monoid and let $U(M)$ denote the group of units of M. If M has rational word problem then $U(M)$ is finite.*

Proof. Let M be a monoid generated by A with $\mathrm{MonWP}(M, A)$ rational and let $C = M \backslash U(M)$.

Note that C is an ideal if and only if every right-invertible element is also left-invertible, if and only if every left-invertible element is right-invertible.

If C is an ideal, then $U(M)$ is finitely generated by $U(M) \cap A$ and has rational word problem by Corollary 4.5. This means that by Theorem 5.1, the group $U(M)$ is finite.

If C is not an ideal, we can pick a from C and and b in $U(M)$ with the property that $ab = 1$ and $ba \neq 1$. By [CP61], Corollary 1.32 the submonoid of M that is generated by a and b is a bicyclic monoid. By Corollary 4.5 and Lemma 3.7 this cannot happen for a monoid with rational word problem. ☐

The following theorem extends Theorem 6.2 to semigroups, stating that every group that is contained in a semigroup with rational word problem has to be finite. This is straightforward for groups that are finitely generated subsemigroups by Theorem 4.3.

{thm:subgrp_rational_fi

Theorem 6.3. *Let S be a semigroup with rational word problem. Then all subsemigroups of S that are groups are finite.*

Proof. Let S be a semigroup finitely generated by A with rational word problem and assume there exists an infinite subsemigroup G of S that is a group. Let $\mathfrak{A}$ be an asynchronous finite state automaton that decides $\mathrm{SgWP}(S, A)$ and let N be the number of states of $\mathfrak{A}$. Let e be the identity of G, let f be a string with $\overline{f} = e$, and let n be the length of f.

Since G is infinite, there exist $g = \overline{w}$ in G with the property that a shortest string w' such that $\overline{wfw'} = e$ has length greater than $(n+1)N + n$.

The automaton accepts (wfw', f), therefore it has to go into a loop while reading a subword of w' on the first and reading nothing on the second tape. This means that there are strings a, b and c with $|b| \geq 1$ such that $w' = abc$ and $(wfab^ic, f)$ is accepted by the automaton for all $i \in \mathbb{N}$, in particular $(wfac, f)$ is accepted by $\mathfrak{A}$. Therefore

$$e = \overline{w}\,\overline{f}\,\overline{ac} = ge\overline{ac}$$

which implies

$$g^{-1} = g^{-1}ge\overline{ac} = ee\overline{ac} = e\overline{ac} = \overline{f}\,\overline{ac} = \overline{fac}$$

in contradiction to the choice of w' as a shortest string such that fw' represents of g^{-1} of this form. ☐

{sec:constructions}

7. CONSTRUCTIONS

In this section we examine natural algebraic constructions or decompositions involving semigroups and show which of them preserve rational word problem. In particular we show that rational word problem is preserved under adding a zero element or an identity and that a semigroup that is a disjoint union of an infinite semigroup and a finite ideal has rational word problem if and only if the infinite semigroup has rational word problem.

{thm:add_zero_one}

Theorem 7.1. *Let S be a finitely generated semigroup. Then the following statements are equivalent.*

{rational}
{add_zero}
{add_one}

(1) *S has rational word problem.*
(2) *S^0 has rational word problem.*
(3) *S^1 has rational word problem.*

Proof. We only prove the equivalence of (1) and (3), the equivalence of (1) and (2) is a special case of Theorem 7.2.

Let $\mathfrak{A} = \langle Q, A, A, q_0, F, \Delta \rangle$ be an asynchronous finite state automaton that decides $\mathrm{SgWP}(S, A)$.

For S^1 we add 1 to the set of generators. To form an automaton that decides $\mathrm{SgWP}(S^1, A \cup \{1\})$ we add transitions $(q, \varepsilon, 1, q)$ and $(q, 1, \varepsilon, q)$ for all $q \in Q$.

If S^1 has rational word problem, we remove 1 from the generating set. By Theorem 4.3, S has rational word problem. ☐

We show that an infinite semigroup that consists of a finite ideal and an infinite semigroup has rational word problem if and only if the infinite semigroup has rational word problem. In particular the equivalence of (1) and (2) in Theorem 7.1 is a special case of Theorem 7.2.

{thm:ideal}

Theorem 7.2. *Let $T = S \cup I$ be a finitely generated semigroup and assume I to be a finite ideal of T and S to be an infinite subsemigroup of T. Then S has rational word problem if and only if T has rational word problem.*

Proof. To show that S has rational word problem if T has rational word problem, let A be a finite generating set for T. The set $B = A \cap S$ generates S and therefore S has rational word problem by Theorem 4.3.

Conversely, let S be finitely generated by B and let $\mathrm{SgWP}(S, B)$ be rational. Denote by l_b for $b \in B$ the map that maps every element i of I to bi and let

$$\varphi_l : B \to \mathrm{T}_I, b \mapsto l_b,$$

where T_I is the full transformation monoid of the set I. We denote concatenation for T_I by $\circ$ for better readability, and $\alpha \circ \beta$ for α and β in T_I means that we first apply β and then α. The map φ_I uniquely extends to a homomorphism φ from A^* to T_I. Also note that since T_I is finite, we may use it as a subset of the set of states in a finite state automaton.

Let $\mathfrak{B} = \langle Q, B, B, q_0, F, \Delta \rangle$ be an asynchronous finite state automaton deciding the word problem for S with respect to the finite generating set B. The idea of the constructed automaton is as follows.

Given two strings v and w over the generating set A of T, to decide whether $\overline{v} = \overline{w}$ we can distinguish the following cases.

(1) None of the two strings contain an element of I and both elements lie in S, or {noneI}
(2) precisely one string contains an element of I, or {oneI}
(3) both strings contain an element of I and both elements lie in I. {bothI}

To construct an automaton that decides the word problem of T we need three components that provide accepting runs for the cases (1) and (3), and for (2) we have to make sure that there is no run that accepts. For (1), we include the automaton $\mathfrak{B}$, for (3) we use a direct product of two copies of T_I that memorises left-transformations of I by S that are read on both tapes and a direct product of two copies of I to compare elements of I.

For a formal construction consider the automaton

$$\mathfrak{A} = \langle R, A, A, r_0, G, \Gamma \rangle ,$$

over the alphabet $A = B \cup I$, with the set

$$R = \{r_0\} \ \cup \ Q \ \cup \ \mathrm{T}_I \times \mathrm{T}_I \ \cup I \times I,$$

of states and the following transition relation in which we denote by α and β elements of T_I, by i and j elements of I, by x and y elements of B and by a and b elements of A,

$$\begin{aligned}
\Gamma \ = \ & \{(r_0, \varepsilon, \varepsilon, q_0)\} \cup \{(r_0, \varepsilon, \varepsilon, (id, id))\} \\
& \cup \Delta \\
& \cup \ \{((\alpha, \beta), x, \varepsilon, (\alpha \circ (\varphi_I x), \beta)) \mid x \in B\} \\
& \cup \ \{((\alpha, \beta), \varepsilon, y, (\alpha, \beta \circ (\varphi_I y))) \mid y \in B\} \\
& \cup \ \{((\alpha, \beta), a, b, (\alpha a, \beta b) \mid a, b \in I\} \\
& \cup \ \{((i, j), a, \varepsilon, (ia, j)) \mid a \in A\} \\
& \cup \ \{((i, j), \varepsilon, b, (i, jb)) \mid b \in A\} .
\end{aligned}$$

The set G of accept states is $F \cup \{(i, i) \mid i \in I\}$.

To prove correctness we show that (v, w) is accepted by $\mathfrak{A}$ if and only if $\overline{v} = \overline{w}$.

Assume that a pair (v, w) is accepted by $\mathfrak{A}$. This means there is an accepting computation γ of $\mathfrak{A}$ on (v, w). If the computation has the form

$$\gamma : r_0 \xrightarrow{(\varepsilon,\varepsilon)} q_0 \xrightarrow{(v,w)}{}^{*} \ q \in F,$$

we are in case (1). By assumption $\mathfrak{B}$ decides $\mathrm{SgWP}(S, B)$, and therefore $\overline{v} = \overline{w}$.

If γ has the form

$$\gamma : r_0 \xrightarrow{(\varepsilon,\varepsilon)} (id, id) \xrightarrow{(v,w)}{}^{*} \ (i, i)$$

for some i in I, we are in case (3) and by construction $\overline{v} = \overline{w}$ because they represent equal elements of I.

Conversely assume $\overline{v} = \overline{w}$ in T. In case (1) there is an accepting computation on $\mathfrak{B}$ by assumption, and thus an accepting computation on $\mathfrak{A}$ exists by construction. In case (3) we can decompose v and w as $v = v_1 a v_2$ and $w = w_1 b w_2$, where v_1 and

w_1 are elements of B^*, a and b are elements of I and v_2 and w_2 are elements of A^*. By construction of $\mathfrak{A}$ the following computation of $\mathfrak{A}$ on (v, w) exists:

$$\gamma : r_0 \xrightarrow{(\varepsilon,\varepsilon)} (id, id) \xrightarrow{(v_1,w_1)} (\varphi_l v_1, \varphi_l w_1) \xrightarrow{(a,b)} ((\varphi_l v_1)a, (\varphi_l w_1)b)$$

$$\xrightarrow{(v_2,w_2)} (((\varphi_l v_1)a)v_2, (((\varphi_l w_1)b)w_2).$$

What is left to show is that $(\varphi_r v_2)(\varphi_l v_1)a = (\varphi_r w_2)(\varphi_l w_1)b$.

$$((\varphi_l v_1)a)v_2 = \overline{v_1 a v_2} = \overline{v_1 a v_2} = \overline{v} = \overline{w} = \overline{w_1 b w_2} = \overline{w_1}\,\overline{b}\,\overline{w_2} = ((\varphi_l w_1)b)w_2$$

□

{sec:products}

8. PRODUCTS

In this section we examine products of semigroups with rational word problem. The direct product of two semigroups with rational word problem does not have rational word problem in general, even if we assume the direct product to be finitely generated. This can most easily be seen by considering $\mathbb{N}_0 \times \mathbb{N}_0$ which does not have rational word problem by 3.6.

It has been shown in [RRW98] that for two finitely generated semigroups S and T the direct product $S \times T$ of S and T is finitely generated if and only if one of the following conditions is true.

(1) S and T are finite,
(2) S is finite and $S^2 = S$,
(3) T is finite and $T^2 = T$, or
(4) $S^2 = S$ and $T^2 = T$.

Following an example which can be found in Remark 7.5 of [RRW98] we consider a finitely generated infinite semigroup S with rational word problem that has the property $S^2 = S$, effectively enabling us to form the finitely generated infinite semigroup $S \times S$.

Example 8.1. *Let S be given by the presentation*

$$S = \mathrm{Sg}\left\langle a, b \mid a^2 = a, ba = b \right\rangle.$$

The semigroup S is infinite, finitely generated and has rational word problem. The direct product S × S is finitely generated but does not have rational word problem.

There is an easily described set of representatives of elements of S consisting of non-empty strings of the form $"a"^{\alpha}"b"^{\beta}$ for $\alpha \in \{0, 1\}$ and $\beta \in \mathbb{N}$.

Consider the automaton $\mathfrak{A}$ depicted in Figure 1. We prove that two non-empty strings v and w over $\{a, b\}$ are accepted by $\mathfrak{A}$ if and only if $\overline{v} = \overline{w}$.

Let v and w be two non-empty strings such that $\overline{v} = \overline{w}$. Then either both begin with a or they both begin with b. In either case the automaton ends up in a final state after reading the first character of both strings. After that, both strings can contain any number of as as long as there is an equal number of bs in both strings. The automaton can just skip occurrences of a until it reaches a b on each tape which it can read then. Now consider $T = S \times S$. Following [RRW98], the resulting semigroup T is finitely generated and finitely presented. A generating set is for example

$$B = \{(a, a), (a, b), (b, a), (b, b)\}.$$

The elements (b^2, b) and (b, b^2) generate a free commutative semigroup of rank 2 in T and therefore by Theorem 4.3 T does not have rational word problem.

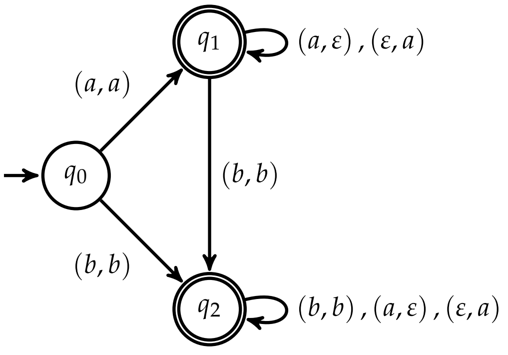


FIGURE 1. Asynchronous finite state automaton $\mathfrak{A}$ that decides $\mathrm{SgWP}(S,\{a,b\})$ for $S = \mathrm{Sg}\left\langle \{a,b\} \mid a^2 = a, ba = b \right\rangle$.

g:directproductexample}

The following theorems characterise how rational word problem behaves under direct products. Given a direct product of two semigroups with rational word problem, it follows that the factors have rational word problem. Conversely, the direct product of two semigroups with rational word problem gives a semigroup with rational word problem only if the direct product is finitely generated and one of the factors is finite.

{thm:dirprod_rat_wp}

Theorem 8.2. *Let S and T be semigroups. If $S \times T$ is finitely generated and has rational word problem, then S and T are finitely generated and have rational word problem.*

Proof. It is sufficient to prove the statement for S. Assume $S \times T$ to be generated by the finite set C. Applying the projection

$$\pi_S : S \times T \to S, (s,t) \mapsto s,$$

to C gives the finite generating set $\pi_S(C)$ for S.

Assume that $S \times T$ has rational word problem and that

$$\mathfrak{A} = \langle Q, A, A, q_0, F, \Delta \rangle$$

is an asynchronous finite state automaton that decides $\mathrm{SgWP}(S \times T, A)$. The following automaton then decides $\mathrm{SgWP}(S, \pi_S(A))$.

$$\mathfrak{A}' = \left\langle Q, \pi_S(A), \pi_S(A), q_0, F, \Delta' \right\rangle,$$

where

$$\Delta' = \{(p, \pi_S a, \pi_S b, q) \mid (p,a,b,q) \in \Delta\}.$$

$\square$

{lem:finite_infinite}

Lemma 8.3. *Let S be a finite semigroup and T be a finitely generated semigroup with rational word problem. If $S \times T$ is finitely generated, then $S \times T$ has rational word problem.*

Proof. Let C be a finite generating set for $S \times T$. We denote by π_S and π_T the projections from $S \times T$ onto S and T respectively.

Since T has rational word problem there is an asynchronous finite state automaton

$$\mathfrak{B} = \langle R, \pi_T(C), \pi_T(C), r_0, G, \Gamma \rangle$$

that decides $\mathrm{SgWP}(T, \pi_T(C))$.

The automaton $\mathfrak{C}$ that decides $\mathrm{SgWP}(S \times T, C)$ can then be given as follows.

$$\mathfrak{C} = \left\langle S^1 \times S^1 \times R, C, C, (1,1,r_0), H, \Pi \right\rangle,$$

where

$$H = \{(s, s, g) \mid s \in S, g \in G\},$$

and the transition relation Π is given as

$$\begin{array}{rcl} \Pi & = & \{[(s,t,q), c, d, (s \cdot \pi_S(c), t \cdot \pi_S(d), r)] \mid (q, \pi_T(c), \pi_T(d), r) \in \Gamma\} \\ & \cup & \{[(s,t,q), \varepsilon, d, (s, t \cdot \pi_S(d), r)] \mid (q, \pi_T(c), \pi_T(d), r) \in \Gamma\} \\ & \cup & \{[(s,t,q), c, \varepsilon, (s \cdot \pi_S(c), t, r)] \mid (q, \pi_T(c), \pi_T(d), r) \in \Gamma\} \\ & \cup & \{[(s,t,q), \varepsilon, \varepsilon, (s, t, r)] \mid (q, \pi_T(c), \pi_T(d), r) \in \Gamma\}. \end{array}$$

We show that (v, w) is accepted by $\mathfrak{C}$ if and only if $\overline{v} = \overline{w}$. Let $\mathfrak{C}$ accept the pair (v, w) in C^+. Then, by construction, there exists an accepting computation on $\mathfrak{B}$, thus $\pi_T(\overline{v}) = \pi_T(\overline{w})$. Also by contruction $\pi_S(\overline{v}) = \pi_S(\overline{w})$.

Now let $\overline{v} = \overline{w}$. In particular $\pi_T(\overline{v}) = \pi_T(\overline{w})$, and hence there exists an accepting run of $\mathfrak{B}$. One can immediately find a run on $\mathfrak{C}$ by lifting this run from $\mathfrak{B}$ to $\mathfrak{C}$. Since also $\pi_S(\overline{v}) = \pi_T(\overline{w})$ the lifted run is accepting. □

{lem:infinite_infinite}

Lemma 8.4. *Let S and T be finitely generated infinite semigroups with rational word problem. Then $S \times T$ contains a free commutative semigroup of rank 2.*

Proof. By Theorem 6.1 there are elements s in S and t in T that generate infinite monogenic subsemigroups in S and T respectively. The elements (s^2, t) and (s, t^2) generate a free commutative semigroup of rank 2 in $S \times T$. Theorem 4.3 now implies that $S \times T$ cannot have rational word problem. □

We summarise the above in the following theorem.

Theorem 8.5. *Let S and T be two semigroups such that $S \times T$ is finitely generated. Then $S \times T$ has rational word problem if and only if at least one of S or T is finite.*

Proof. If $S \times T$ has rational word problem, then Theorem 8.2 implies that S and T have rational word problem.

Conversely, if both S and T are finite then the direct product $S \times T$ is finite and therefore has rational word problem. If S is finite and T is infinite or vice versa, we use Lemma 8.3. Finally, if both S and T are infinite Lemma 8.4 proves that $S \times T$ does not have rational word problem. □

Inductively it follows that any finite direct product $S_1 \times \cdots \times S_n$ of semigroups has rational word problem if and only if it is finitely generated and there is at most one S_i that is infinite and has rational word problem.

Rational word problem is not preserved under monoid free products. Consider the cyclic group C_2. The monoid free product of two copies of C_2 is an infinite group. But infinite groups do not have rational word problem by Theorem 5.1. Thus monoid free products even of finite monoids with rational word problem do not necessarily have rational word problem. The situation is different for semigroup free products.

Theorem 8.6. *Let S and T be two semigroups generated by finite sets A and B respectively. The semigroup free product $S \star T$ has rational word problem if and only if S and T have rational word problem.*

Proof. Let S and T be semigroups with rational word problem and let $\mathfrak{A}$ be an asynchronous automaton that decides $\mathrm{SgWP}(S, A)$, and $\mathfrak{B}$ be an asynchronous automaton that decides $\mathrm{SgWP}(T, B)$. An automaton that decides $\mathrm{SgWP}(S * T, A \cup B)$ can be constructed by using both $\mathfrak{A}$ and $\mathfrak{B}$ and adding a new initial state q_0 and

$(\varepsilon, \varepsilon)$ transitions from q_0 to the initial states of $\mathfrak{A}$ and $\mathfrak{B}$ as well as from the accept states of both automata to q_0.

The converse follows directly from Theorem 4.3. $\square$

Another product construction that is possible for semigroups is the zero union of two semigroups. We define the zero union as follows.

Definition 8.7. *Let U be a semigroup with zero. If there exist subsemigroups S and T of U such that $S \cap T = \{\}$ and $T = S \cup U$ and $st = 0 = ts$ for all $s \in S$ and $t \in T$ then U is a* zero union *of S and T, denoted by $S \cup_0 T$.*

Note that $S \cup_0 T$ is finitely generated if and only if S and T are finitely generated. A generating set for S can be obtained from a generating set C for $S \cup_0 T$ by the intersecting C with S, a generating set for T can be obtained by intersecting C with T. Given generating sets for S and T the union of those generating sets together with the zero element gives a generating set for $S \cup_0 T$. Rational word problem is preserved under zero union.

Theorem 8.8. *Let U be a finitely generated semigroup that is a zero union of two subsemigroups S and T. Then U has rational word problem if and only if S and T have rational word problem.*

Proof. If $U = S \cup_0 T$ has rational word problem, then S and T are finitely generated subsemigroups of U and therefore have rational word problem by Theorem 4.3.

Conversely let C be a generating set for U. Let $A = C \cap S$ and let $B = C \cap T$ be generating sets for S and T respectively and assume that $\operatorname{SgWP}(S, A)$ and $\operatorname{SgWP}(T, B)$ are rational.

Additionally we observe that the set

$$Z = \left\{ v \in C^+ \mid \overline{v} = 0 \right\},$$

is regular and hence $Z \times Z$ is rational. We show that

$$\operatorname{SgWP}(S \cup_0 T, C) = \operatorname{SgWP}(S, A) \cup \operatorname{SgWP}(T, B) \cup (Z \times Z).$$

Let (v, w) be in $\operatorname{SgWP}(S \cup_0 T, C)$, which is the case if and only if $\overline{v} = \overline{w}$ and we distinguish three cases

1. $\overline{v}$ is a non-zero element of S,
2. $\overline{v}$ is a non-zero element of T, or
3. $\overline{v}$ is zero.

In the first two cases (v, w) is contained in the right hand side, because it is either contained in $\operatorname{SgWP}(S, A)$ or in $\operatorname{SgWP}(T, B)$ respectively. A string v over C represents the zero element of $S \cup_0 T$ if and only if it is contained in Z, thus if $\overline{v} = 0$ then (v, w) is contained in $Z \times Z$. $\square$

9. REMARKS AND OUTLOOK

`{sec:remarks_and_outloo`

This section aims at giving reference to further research questions which are outside the scope of this paper.

Obviously a semigroup with rational word problem has decidable word problem and it is undecidable whether a given finitely generated semigroup has rational word problem because then it would be decidable whether a given finitely generated semigroup was trivial, because it can be easily checked whether a semigroup with rational word problem is trivial. One interesting project is to find an algorithm that, given a presentation of a monoid finds an automaton that decides the word problem if it exists.

Once one has characterised all semigroups with rational word problem, one also has classified all rational congruences. This is because the word problem of a semigroup S finitely generated by a subset A is the kernel of the canonical map $^{-} : A^{+} \to S$, and every rational congruence is the kernel of such a map. An open question that is tied to this is whether rational equivalence relations have regular cross sections, that is can we find a regular set of unique representatives for each equivalence class or even a finite state automaton that computes for any given word this normal form. This problem was investigated in [Joh85] and to this day was not solved.

Further research will aim at finding a full characterisation of all semigroups with rational word problem, extending the notion of rational word problem to intersections of rational relations as well as giving a better picture of the relationship between different definitions of word problem and different automaton models that decide those word problems. This will provide a more complete picture of the complexity of word problems that arise. In connection to the structure theory of semigroups, there are questions to be asked about Green's relations. For example are Green's relations rational for semigroups with rational word problem? How many $\mathcal{R}-$ or $\mathcal{L}-$classes can a semigroup with rational word problem have? Are all $\mathcal{H}$ classes of such semigroups finite?

A further direction of research is connections to geometric semigroup theory because there is an interesting connection between the theory of automatic groups and geometric group theory as described in [EPC^{+}92] it should be determined in how far there might be a connection of this type for the theory presented in this paper.

10. CONCLUSION

We have introduced a natural class of finitely generated semigroups with the property that the word problem is decidable by an asynchronous finite state automaton. We showed this property to be independent of the generating set and then also showed behaviour of the property under a few basic constructions. Also it was shown that all finite groups are contained in this class but no infinite ones. There are some very simple infinite semigroups with rational word problem.

We were not yet able to achieve a full characterisation of all semigroups with rational word problem. This is due to the fact that there is no nice decomposition theory for semigroups, but this goal seems to be achieveable. It is also desirable to develop this theory with less ad-hoc proofs for the automata theoretic theorems.

REFERENCES

[Ani71] A. V. Anisimov, *On group languages*, Kibernetika **4** (1971), 18–24.

[Ber79] Jean Berstel, *Transductions and context-free languages*, Teubner Studienbücher, Stuttgart, 1979.

[CP61] A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups*, vol. 1, American Mathematical Society, 1961.

[CP67] ———, *The algebraic theory of semigroups*, vol. 2, American Mathematical Society, 1967.

[EPC^{+}92] David B. A. Epstein, M. S. Paterson, G. W. Camon, D. F. Holt, S. V. Levy, and W. P. Thurston, *Word processing in groups*, A. K. Peters, Ltd., Natick, MA, USA, 1992.

[HMU01] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman, *Introduction to automata theory, languages, and computation*, 2nd ed., Addison-Wesley, Boston, 2001, 00064608 John E. Hopcroft, Rajeev Motwani, Jeffrey D. Ullman. Includes bibliographical references (p. 510-511) and index.

[HOT08] D.F. Holt, M.D. Owens, and R.M. Thomas, *Groups and semigroups with a one-counter word problem*, Journal of the Australian Mathematical Society **85** (2008), 197–209.

[Joh83] John Howard Johnson, *Formal models for string similarity*, Ph.D. thesis, University of Waterloo, School of Computer Science, 1983, AAI0553315.

[Joh85] J. Howard Johnson, *Do rational equivalence relations have regular cross-sections?*, Proceedings of the 12th Colloquium on Automata, Languages and Programming (London, UK), Springer-Verlag, 1985, pp. 300–309.

[Kam09a] Mark Kambites, *Small overlap monoids i: The word problem*, Journal of Algebra **321** (2009), no. 8, 2187 – 2205, Computational Algebra.

[Kam09b] ———, *Small overlap monoids ii: Automatic structures and normal forms*, Journal of Algebra **321** (2009), no. 8, 2302 – 2316, Computational Algebra.

[RRW98] E.F. Robertson, N. Ruskuc, and J. Wiegold, *Generators and relations of direct products of semigroups*, Transactions of the American Mathematical Society **350** (1998), no. 7, 2665 – 2685.

E-mail address: `neunhoef@mcs.st-and.ac.uk`

E-mail address: `markusp@mcs.st-and.ac.uk`

E-mail address: `nik@mcs.st-and.ac.uk`

SCHOOL OF MATHEMATICS AND STATISTICS, UNIVERSITY OF ST ANDREWS, MATHEMATICAL INSTITUTE, NORTH HAUGH, ST ANDREWS, KY16 9SS, SCOTLAND