Word problems and formal language theory

Rick Thomas

Department of Computer Science

University of Leicester

CSA 2016

20th June 2016

Happy birthday to Jorge and Gracinda!

Generating sets

If Σ is a finite set of symbols we let Σ^* denote the set of all finite words of symbols from Σ (including the empty word ϵ). If we only want to consider non-empty words we denote the resulting set by Σ^+.

Σ^+ is the *free semigroup* on Σ and Σ^* is the *free monoid* on Σ.

If we have a group G (or a monoid M) with a finite set of generators Σ, then we have a natural homomorphism $\varphi : \Sigma^* \longrightarrow G$ (or $\varphi : \Sigma^* \longrightarrow M$).

For a semigroup S generated by a finite set Σ we have a natural homomorphism $\varphi : \Sigma^+ \longrightarrow S$.

Word problems

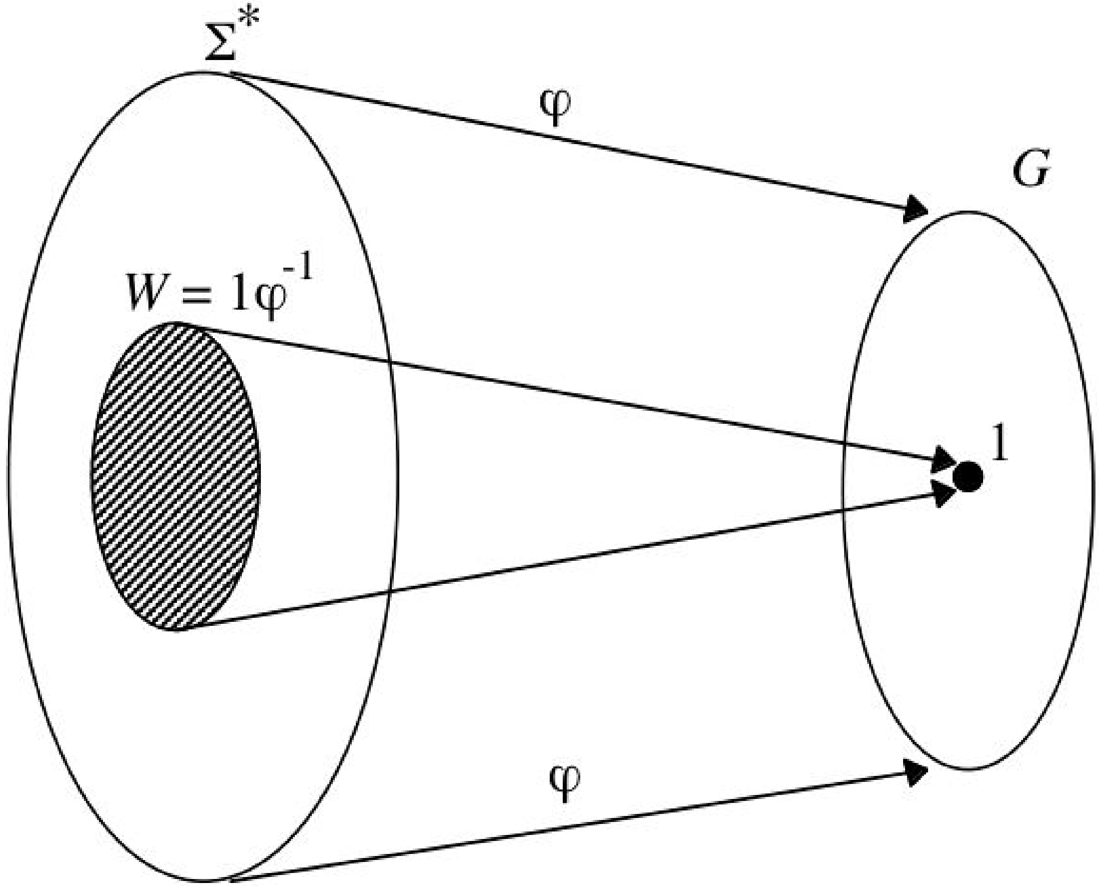
The *word problem* in such a structure is the following question:

Input: Two words α and β in Σ^* (or Σ^+ in the case of a semigroup);

Output: **Yes** if α and β represent the same element of the group (monoid, semigroup);

No otherwise.

In a group, given a word β representing an element g, let γ be a word representing g^{-1}. Now α and β represent the same element of the group if and only if $\alpha\gamma$ represents the identity.



Σ^*
φ
G
$W = 1\varphi^{-1}$
1
φ

Word problems

Given this, we can define the *word problem* $W = W(G)$ of a finitely generated group G to be the set of all words in Σ^* that represent the identity element of G. (This is not appropriate for monoids and does not make sense in semigroups.)

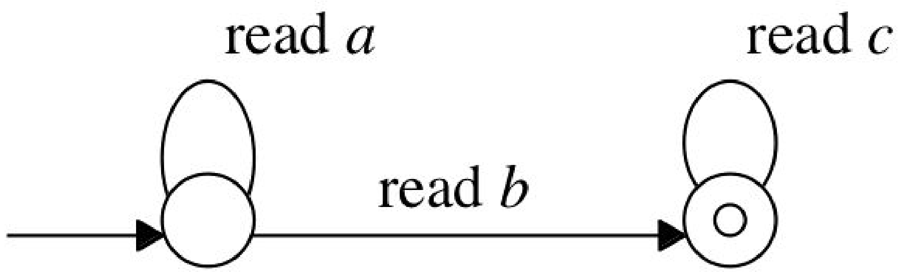
In this way, we can think of the word problem of a group as being a formal language.

We will focus on some relatively simple classes of languages, the *regular languages*, the *one-counter languages* and the *context-free languages*. Saying that the word problem of a group is regular (or one-counter or context-free) does not depend on the choice of finite generating set.

Automata

We can define classes of languages using various notions of "automata".

Regular languages are accepted by *finite automata*.

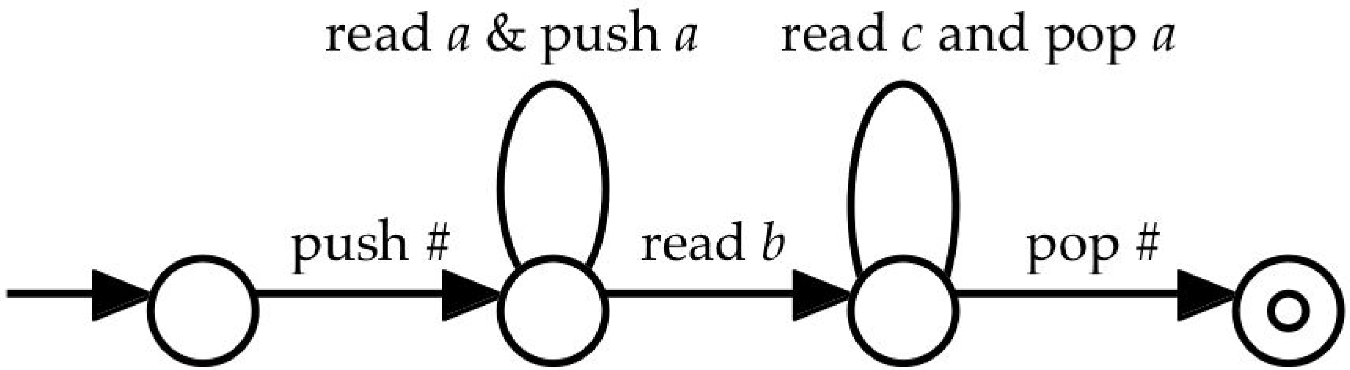


A word is *accepted* if we reach an "accept state" after reading the word.

The *language* $\mathrm{L}(\mathrm{M})$ of M is the set of all words accepted by M.

Pushdown automata

Context-free languages are accepted by *pushdown automata* where we add a "stack" to the machine.



If we restrict to one stack symbol (apart from a fixed bottom marker #) we have a *one-counter language*.

An automaton is said to be *deterministic* if there can never be a possibility of choice as regards to which move to make.

Word problems of groups

If G is a finitely generated group, then $W(G)$ is regular if and only if G is finite. (Anisimov)

If G is a finitely generated group, then $W(G)$ is context-free if and only if G is virtually free. (Muller & Schupp)

As a consequence, if $W(G)$ is context-free, then it is deterministic context-free.

If G is a finitely generated group, then $W(G)$ is a one-counter language if and only if G is virtually cyclic. (Herbst)

Word problems of groups

The following are equivalent for a finitely generated group G and $n \geqslant 1$:

(i) The word problem of G is the intersection of n one-counter languages.

(ii) The word problem of G is the intersection of n deterministic one-counter languages.

(iii) G is virtually abelian of free abelian rank $\leqslant n$. (Holt, Owens & Thomas)

G being virtually abelian is also equivalent to the word problem of G being a Petri net language. (Rino Nesin & Thomas)

Conjecture. The word problem of a finitely generated group G is the intersection of n context-free languages (for some n) if and only if G is virtually a finitely generated subgroup of a direct product of free groups. (Brough)

Word problems of groups

A language L over an alphabet Σ is the word problem of a group with generating set Σ if and only if L satisfies the following two conditions:

(W1) for all $\alpha \in \Sigma^*$ there exists $\beta \in \Sigma^*$ such that $\alpha\beta \in L$;

(W2) if $\alpha\delta\beta \in L$ and $\delta \in L$ then $\alpha\beta \in L$. (Parkes & Thomas)

As a consequence of the Muller-Schupp clasification we have:

If L is a context-free language satisfying (W1) and (W2) then L is deterministic context-free.

There are many other such classifications and associated decidability results. (Jones & Thomas)

Decidability

There is no algorithm that, given a context-free language L, will decide whether or not L is the word problem of a group. (Lakin & Thomas)

This can be generalized to the fact that there is no algorithm that, given a one-counter language L, will decide whether or not L is the word problem of a group. (Jones & Thomas)

However, there is an algorithm that, given a deterministic context-free language L, will decide whether or not L is the word problem of a group. (Jones & Thomas)

Word problems of semigroups

Duncan and Gilman proposed the following definition of the word problem for a semigroup S generated by a finite set A:

$$W(S) = \{\alpha\#\beta^{rev} : \alpha, \beta \in A^+, \alpha =_S \beta\}.$$

This is a natural generalization of the word problem of a group G which was

$$W(G) = \{\alpha\beta^{-1} : \alpha, \beta \in A^*, \alpha =_G \beta\}.$$

In this way, we can consider the word problem of a semigroup as a formal language.

If S is a finitely generated semigroup, then $W(S)$ is regular if and only if S is finite. (Duncan & Gilman)

One-counter word problems

If a finitely generated semigroup S has word problem a one-counter language, then S has a linear growth function. (Holt, Owens & Thomas)

If S is a finitely generated semigroup with a linear growth function then there exist finitely many elements $a_i, b_i, c_i \in S \cup \{\epsilon\}$ such that every element of S is represented by a word of the form $a_i b_i^n c_i$ for some i and some $n \geqslant 0$. (Holt, Owens & Thomas)

For context-free word problems in semigroups there are some partial results (Hoffmann, Holt, Owens & Thomas) but we are far from a classification.

Thank you!