Formal Languages and Groups

Rick Thomas

University of Leicester

Department of Computer Science

http://www.cs.le.ac.uk/~rthomas

Antalya Algebra Days VII, 22nd May 2005

Σ : finite set of symbols.

Σ^* : the set of all finite words formed from the symbols in Σ (including the *empty word* ε).

Example.

$\Sigma = \{ a, b \}$;

$\Sigma^* = \{ \varepsilon, a, b, a^2, ab, ba, b^2, a^3, a^2b, \ldots\ldots\ldots \}$.

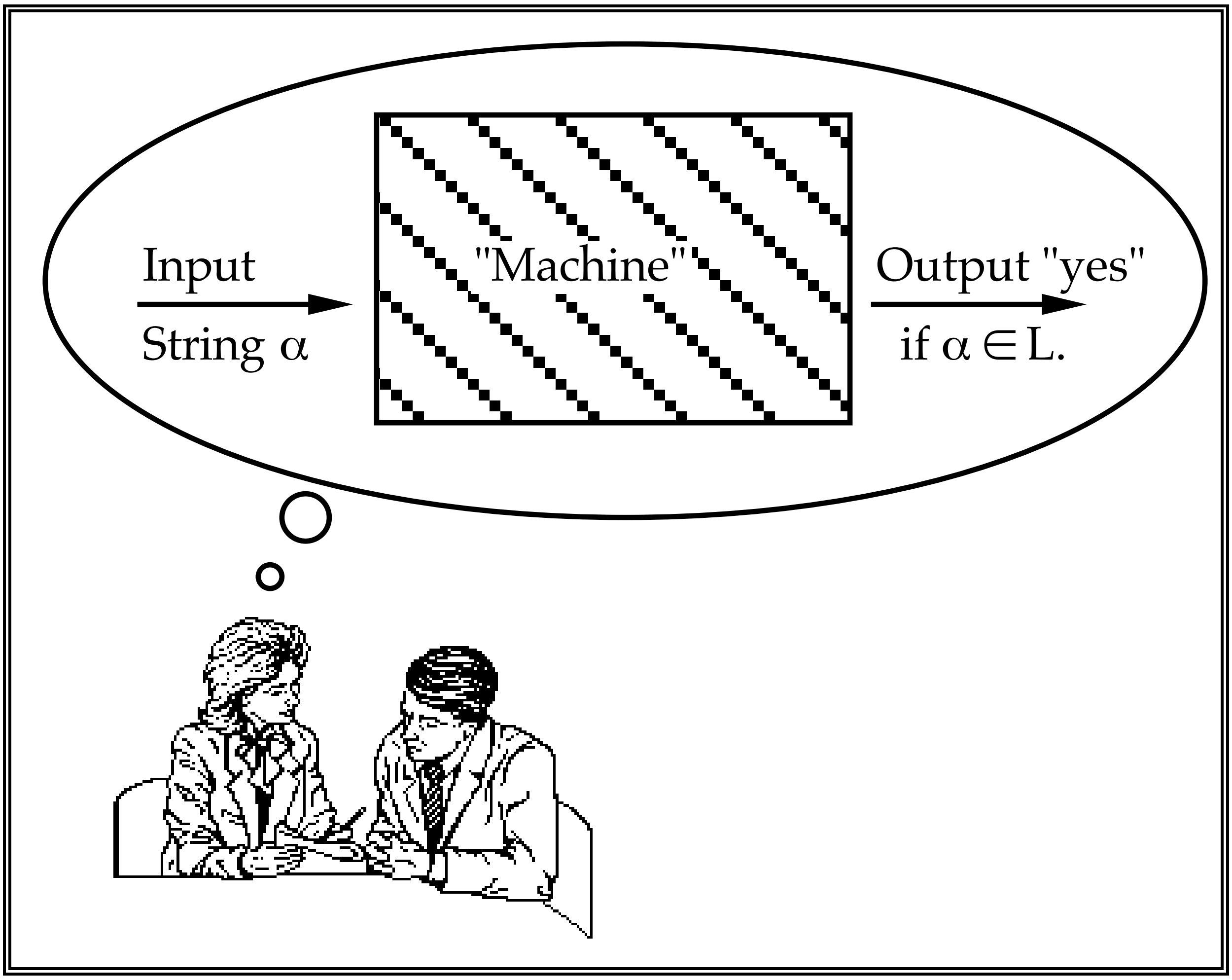
A *language* is a subset of Σ^* (for some finite set Σ).

Example.

$\Sigma = \{ a, b \}$;

$L = \{ \alpha \in \Sigma^* : |\alpha| \text{ is even}\}$.

One can define a hierarchy of abstract models of machines accepting a range of families of languages.

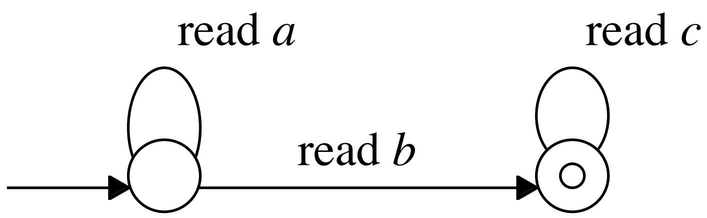


<table>
<tr><td>Regular languages</td><td>Finite automata</td></tr>
<tr><td>Context-free languages</td><td>Pushdown automata</td></tr>
<tr><td>Recursive languages</td><td rowspan="2">Turing machines</td></tr>
<tr><td>Recursively enumerable languages</td></tr>
</table>

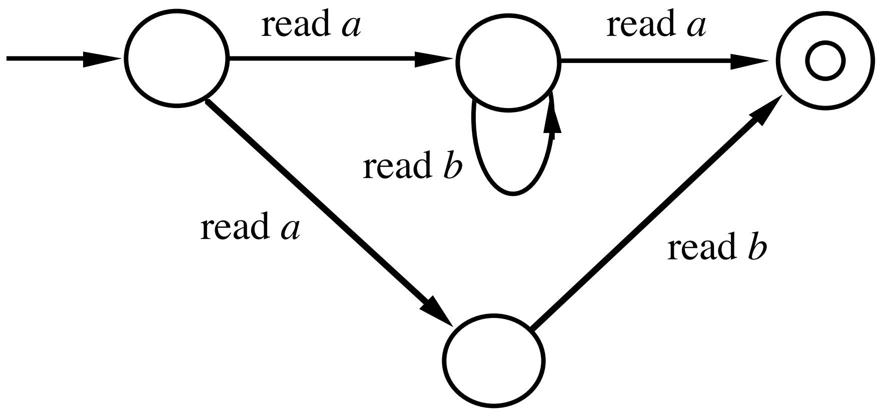
Regular languages

Regular languages are accepted by *finite automata*.

The finite automaton below accepts the language $\{a^n b c^m : n, m \in \mathbf{N}\}$:



Allowing non-determinism, such as



does not increase the range of languages accepted.

In a non-deterministic machine, a word is accepted if at least one computation path leads to acceptance; in our example, the machine accepts

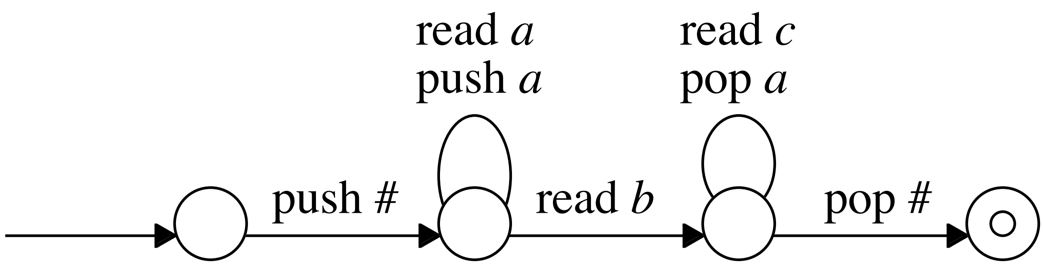
$$\{ab^n a : n \in \mathbf{N}\} \cup \{ \, ab \, \}.$$

Context-free languages

We may extend a (non-deterministic) finite automaton by adding a *stack* to get a *pushdown automaton*.

The languages accepted by pushdown automata are known as *context-free languages*.

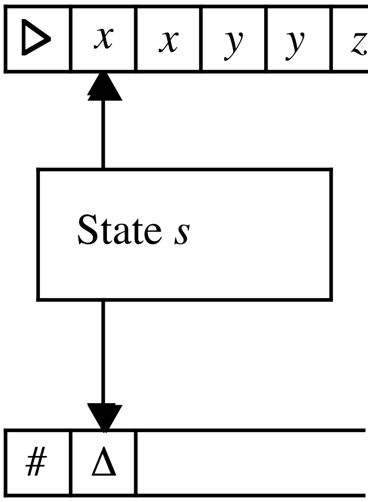
The pushdown automaton below accepts the language $\{a^n b c^n : n \in \mathbf{N}\}$:



Insisting that the machine is deterministic does restrict the range of languages accepted in this case.

Recursive languages

A most general model of computation is the *Turing machine*; here we have a *halt state*.

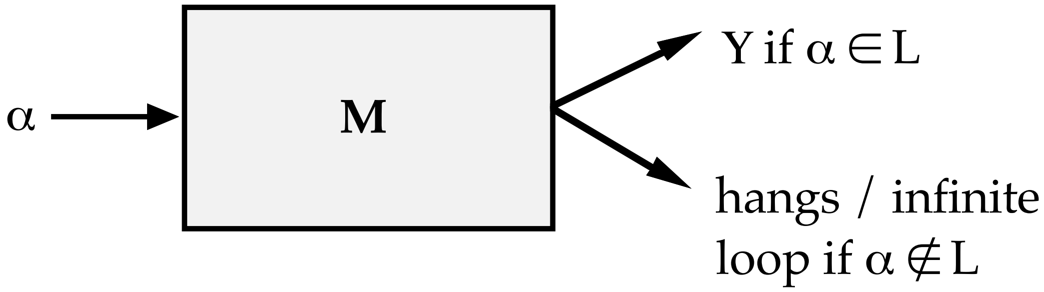


A Turing machine with a given input α will either

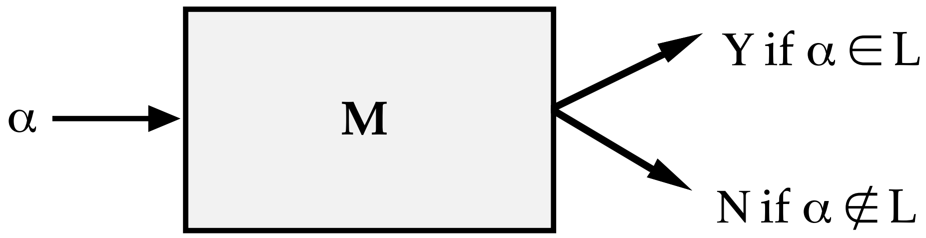
(i) terminate (if it enters the halt state), in which case α is accepted; or

(ii) hang or run indefinitely ("infinite loop").

Recursively enumerable



Recursive



Allowing non-determinism does not increase the range of languages accepted or decided by Turing machines.

However, if we put restrictions on the amount of time or space allowed for the computation, the situation may change.

P = NP Question.

Is the class **P** of languages accepted by deterministic Turing machines in polynomial time equal to the class **NP** of languages accepted by non-deterministic Turing machines in polynomial time ?

Groups

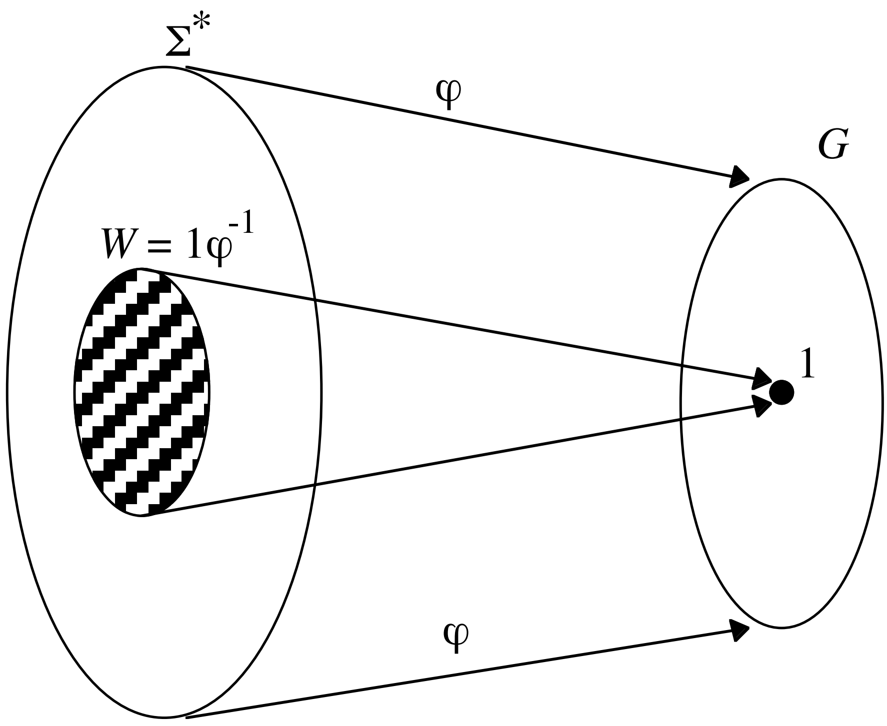
Let G be a group generated by a finite set X so that each element of G may be written as a word in the symbols x, x^{-1} (where $x \in X$). Let $\Sigma = X \cup X^{-1}$.

A group G is said to be *finitely presented* if it has a presentation $< X : R >$ with both X and R finite.

Word problem for groups.

Given a presentation $< X : R >$ defining a group G and given $\alpha \in \Sigma^*$, is $[\alpha] = 1$ in G ?

Alternatively, we can define the *word problem* W for a group G generated by a finite set X to be the set of all words in Σ^* that represent the identity element of G.



What does the nature of the word problem say about the algebraic structure of G ?

If G has a finite presentation $< X : R >$ then $W_X(G)$ is recursively enumerable.

The word problem for G is said to be *solvable* if $W_X(G)$ is recursive.

Novikov, Boone. There exist finitely-presented groups with insolvable word problem.

Higman. If G is a group generated by a finite set X, then $W_X(G)$ is recursively enumerable if and only if G is a subgroup of a finitely presented group.

Boone & Higman. $W_X(G)$ is recursive if and only if $G \leq H \leq K$, where H is simple and K is finitely presented.

Birget, Ol'shanskii, Rips & Sapir.
There is a finitely presented group whose word problem is **NP** -complete.

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

Example.

$X = \{a, b\}$; G free group on X.

$\Sigma = \{ a, a^{-1}, b, b^{-1} \}$.

$W(G)$ is the minimal language L such that:

(i) $\varepsilon \in L$;

(ii) $\alpha \in L \Rightarrow a\alpha a^{-1},\ b\alpha b^{-1},\ a^{-1}\alpha a,\ b^{-1}\alpha b \in L$;

(iii) $\alpha, \beta \in L \Rightarrow \alpha\beta \in L$.

$W(G)$ is context-free but not regular.

If G and H are groups with presentations $< X : R >$ and $< Y : S >$ and subgroups K_1 and K_2 respectively, together with an isomorphism $\varphi : K_1 \to K_2$, we write $G *_K H$ for the *free product* of G and H *with K amalgamated*, i.e. the group with presentation

$$< X \cup Y : R \cup S \cup \{ k = k\varphi : k \in K_1 \} >.$$

If $K = \{ 1 \}$, we just write $G * H$.

A (finitely generated) *free group* is then a group of the form $\mathbf{Z} * \mathbf{Z} * \ldots * \mathbf{Z}$.

Muller & Schupp (and Dunwoody). If G is a finitely generated group, then $W(G)$ is context-free if and only if G has a free subgroup of finite index.

Example.

$G = < a, b : ab = ba > = \mathbf{Z} \times \mathbf{Z}$.

$\Sigma = \{ a, a^{-1}, b, b^{-1} \}$.

For any word α, let $\sigma_a(\alpha)$ and $\sigma_b(\alpha)$ denote the exponent sums of a and b in α respectively.

$$W(G) = \{ \alpha \in \Sigma^* : \sigma_a(\alpha) = \sigma_b(\alpha) = 0 \}.$$

$W(G)$ is not context-free.

However, Σ^* - $W(G)$ is context-free.

Question. Which finitely-generated groups have a word problem which is the complement of a context-free language?

Let $\mathcal{C}$ denote the class of such groups.

Some facts (Holt/Rees/Röver/Thomas).

- $\mathcal{C}$ contains all free and abelian groups.
- $\mathcal{C}$ is closed under taking finitely generated subgroups.
- $\mathcal{C}$ is closed under taking finite index overgroups.
- $\mathcal{C}$ is closed under direct products.
- $\mathcal{C}$ is closed under taking wreath products with virtually free groups.
- there exist groups in $\mathcal{C}$ that are not finitely presented.

Question.

- Is $\mathcal{C}$ closed under free products?

Reduced and irreducible word problems

If $G = < X >$ and $W = W_X(G)$, we define the *reduced word problem* to be

$$R_X(G) = \{ \alpha \in W : \alpha \neq \varepsilon \text{ and no non-empty proper prefix of } \alpha \text{ is in } W \},$$

and the *irreducible word problem* to be

$$I_X(G) = \{ \alpha \in W : \alpha \neq \varepsilon \text{ and no non-empty proper subword of } \alpha \text{ is in } W \}.$$

A language L is a *simple language* if it is accepted by a one-state deterministic pushdown automaton that accepts by empty stack.

Haring-Smith. If G is a finitely generated group, then the following are equivalent:

(i) $R_X(G)$ is a simple language for some finite group generating set X.

(ii) $I_X(G)$ is finite for some finite group generating set X.

(iii) G is of the form $F*G_1*G_2*....*G_r$ where F is free and each G_i is finite.

Such a group is called a *plain group*.

Conjecture. The groups with deterministic context-free reduced word problem are precisely the finite extensions of plain groups.

Parkes & Thomas. Let G be a group generated by a finite set X. Then:

$R_X(G)$ is context-free

$\Leftrightarrow$ $R_X(G)$ is deterministic context-free

$\Leftrightarrow$ $W_X(G)$ is context-free.

Madlener & Otto. Which groups have finite irreducible word problem with respect to some monoid generating set?

$G = \langle f : \ \rangle \times \langle a : a^2 = 1 \rangle = \mathbf{Z} \times \mathbf{Z}_2$.

$X = \{ f, g \}$, $g = af^{-1}$ - a finite monoid generating set for G.

G is not a plain group. $I_X(G)$ is finite.

Some more non-plain groups with finite irreducible word problem:

1. $\mathbf{Z} \times \mathbf{Z}_k$ $(k \geq 2)$. 2. $\mathbf{Z}_{2k} *_{\mathbf{Z}_k} \mathbf{Z}_{2k}$ $(k \geq 2)$.

What happens with regular or context-free?

$I_X(G)$ regular $\Rightarrow$ $I_X(G)$ finite.

$I_X(G)$ context-free $\Rightarrow$ $W_X(G)$ context-free.

Family F of languages	Groups G with $I_X(G) \in$ F (X a group generating set)
Finite	Plain groups
Regular	Plain groups
Context-free	?