The Word Problem for Z^2

Robert Gilman
Stevens Institute of Technology

Equations and Formal Languages in Algebra

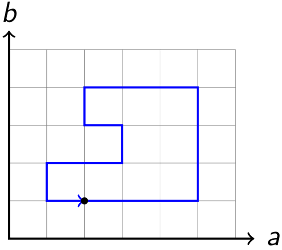
Les Diablerets
March 9, 2016

Revised for online viewing.

The word problem

G is a finitely presented group, e.g. $G = \langle a, b \mid aba^{-1}b^{-1} \rangle = Z^2$.

The word problem: decide when a word w in the generators and their inverses represents the identity in G, i.e., when w is the label of a cycle in the Cayley diagram Γ of G.



A cycle in Γ with label $a^3b^3a^{-3}b^{-1}ab^{-1}a^{-2}b^{-1}a$.

W is the formal language of labels of cycles in Γ.

The word problem has algebraic, geometric, and language-theoretic aspects.

Anisimov and co-workers in Kiev were the first to view word problems as formal languages.

Theorem [Anisimov, 1971] G is finite if and only if W is regular.

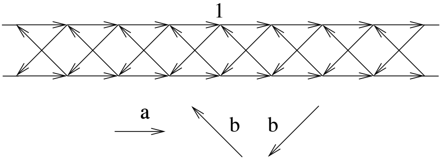
Theorem [Muller and Schupp, 1981] The following are equivalent.

- The word problem of G is context free.
- G has a free subgroup of finite index.
- The Cayley diagram of G has finitely many end-isomorphism types.
- Every cycle in the Cayley diagram of G is triangulated by chords of uniformly bounded length.

A beautiful combination of algebra, language theory and geometry.

For a nice new proof see Diekert and Weiss, 2013.

A Cayley diagram with finitely many end isomorphism types



Removing the ball of radius k around 1 yields only finitely many isomorphism types of connected components as k goes to infinity.

Dependence of W on choice of generators

Every finite presentation $G = \langle \Sigma \mid R \rangle$ determines a surjective morphism $\pi : (\Sigma \cup \Sigma^{-1})^* \to G$ with $W = \pi^{-1}(1)$.

Context free languages are closed under inverse homomorphism; so if W is context free for one presentation of G, then it is for all.

$$\begin{tikzcd}
D \arrow[d, "s"'] \arrow[dr, dotted, "f"] & \\
G & S \arrow[l, "p"']
\end{tikzcd}$$

The same argument shows that if G has context free word problem, then every finitely generated subgroup of G does too.

Some recent developments concerning word problems

(Word problems have been studied since 1911.)

Context free co-word problems [Holt, Rees, Röver] and [Lehnert, Schweitzer]. The complement of W is context-free.

Poly-context-free word problems [Brough]. W is an intersection of context-free languages. Conjectured to characterize finite extensions of finitely generated subgroups of direct products of free groups.

Multipass word problems. [Ceccherini-Silberstein, Coornaert, Fiorenzi, Schupp, Touikan]. W is in the Boolean closure of deterministic context-free languages. Also conjectured to characterize finite extensions of finitely generated subgroups of direct products of free groups.

More general language-theoretic structures for groups

The general idea is to consider a language L which maps surjectively to G and impose conditions on some language related to the word problem with regard to L.

Automatic groups.

Word hyperbolic groups.

Graph automatic groups.

$\mathcal{C}$-Graph automatic groups [Elder, Taback 2014]. Includes Baumslag-Solitar groups.

Two questions which have been open for a long time

Question 1. Is there a group whose word problem is indexed but not context-free?

Question 2. Is Z^2 such a group?

Recently Sylvain Salvati cast some new light on these questions by proving that the word problem of Z^2 is a multiple context-free language (MCFL).

MCFL's originated in mathematical linguistics as a model for natural languages. Context free languages are not an adequate model, because of cross serial dependencies [1985].

dass de Karl d'Maria em Peter de Hans laat hälfe lärne schwüme

daß der Karl die Maria dem Peter den Hans schwimmen lehren helfen läßt

that Charles lets Mary help Peter teach John to swim.

The following sentence is grammatical if and only if $k = m$ and $l = n$.

De Jan säit, dass mer (d'chind)^k (em Hans)^l es huus haend wele laa^m hälfe^n aastriiche.

It follows that there is a transduction from Swiss German to the language $\{wa^i b^j xc^i d^j y\}$. As this lanuage is not context-free, Swiss German is not context free either.

Mild context sensitivity is required for an adequate model of natural language.

- Limited cross-serial dependencies.
- Constant growth property.
- Polynomial time parsing.

Multiple context-free grammars (MCFG's)

A context-free grammar:

Alphabet: $\{a, b, \$\}$

Productions: $S \to \$A\$$, $A \to ab$, $A \to aAb$

Sample derivation: $S \to \$A\$ \to \$aAb\$ \to \$aabb\$$

Language generated: $\{\$a^n b^n \$ \mid n \geq 1\}$

In MCFG's nonterminals can derive (i.e., yield) tuples of words.

Alphabet: $\{a, b, c, d, \$\}$

Rules:

$A[ab, cd] \leftarrow$	A yields the pair of words (ab, cd)
$S[\$xy\$] \leftarrow A[x, y]$	If A yields (x, y), then S yields $\$xy\$$
$A[axb, cyd] \leftarrow A[x, y]$	If A yields (x, y), then A yields (axb, cyd)

Sample derivation:

Rule	$S \leftarrow A$	$A \leftarrow A$	$A \leftarrow$
Yield	$\$aabbccdd\$$	$(aabb, ccdd)$	(ab, cd)

Language generated: $\{\$a^n b^n c^n d^n \$ \mid n \geq 1\}$

The general form of a rule over the alphabet Σ is

$$A_0[t_1, \dots, t_m] \leftarrow A_1[x_{1,1}, \dots, x_{1,i_1}], \dots, A_n[x_{n,1}, \dots, x_{n,i_n}]$$

1. The $x_{i,j}$'s are variables.
2. The t_i's are words over $\Sigma \cup \{x_{i,j}\}$;
3. The A_j's are nonterminal symbols.
4. The A_j's may have repetitions, but the $x_{i,j}$'s are distinct.
5. Each variable $x_{i,j}$ occurs at most once in the word $t_1 t_2 \cdots t_m$.

m is the rank of A_0 and $i_1, \dots, i_n$ are the ranks of $A_1, \dots, A_n$ respectively.

The grammar is a k-MCFG if the maximum rank is k. The rank of S is always 1.

k-MFCL's form a strictly increasing sequence of full abstract substitution-closed families of languages.

1-MFCL's are context free languages.

Every infinite k-MCFL, L, satisfies a weak form of the pumping lemma.

There exists $w = u_0 v_1 u_1 v_2 \cdots v_{2k} u_{2k} \in L$ with $\sum |v_j| > 0$ and $w = u_0 v_1^i u_1 v_2^i \cdots v_{2k}^i U_{2k} \in L$ for all $i \geq 0$.

Theorem. Let $\mathcal{C}(k)$ be the collection of groups with a word problem W which is a k-MCFL.

- Membership in $\mathcal{C}(k)$ is independent of choice of generators.
- $\mathcal{C}(k)$ is closed under finitely generated subgroup and finite index supergroup.

Another MCFG

The alphabet is $\Sigma = \{a, b, c, d\}$. ε denotes the empty word.

$S[x_1 y_1 x_2 y_2] \leftarrow P[x_1, x_2], Q[y_1, y_2]$
$P[ax_1, dx_2] \leftarrow P[x_1, x_2]$
$P[\varepsilon, \varepsilon] \leftarrow$
$Q[cx_1, dx_2] \leftarrow Q[x_1, x_2]$
$Q[\varepsilon, \varepsilon] \leftarrow$

The language generated is $\{a^n c^m b^n d^m \mid m, n \geq 0\}$.

The first rule says that if (x_1, x_2) is in the yield of P and (y_1, y_2) is in the yield of Q, then $x_1 y_1 x_2 y_2$ is in the yield of S.

Or if we think of the nonterminals as predicates over Σ^*, then the first rule says that if $P[x_1, x_2]$ and $Q[y_1, y_2]$ are true, so is $S[x_1 y_1 x_2 y_2]$.

The word problem of Z^2

$Z^2 = \langle a, b \mid aba^{-1}b^{-1} \rangle$

Theorem 1. [Salvati] The following multiple context-free grammar generates the word problem for Z^2.

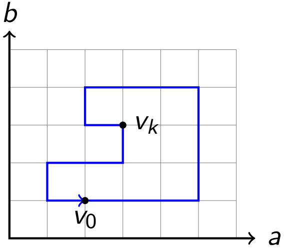
1. $S[x_1 x_2] \leftarrow A[x_1, x_2]$
2. $A[t_1, t_2] \leftarrow A[x_1, x_2]$ where the word $t_1 t_2$ is any permutation of $\{x_1, x_2, a, a^{-1}\}$ or $\{x_1, x_2, b, b^{-1}\}$
3. $A[t_2, t_2] \leftarrow A[x_1, x_2], A[y_1, y_1]$ where $t_1 t_2$ is any permutation of $\{x_1, x_2, y_1, y_2\}$
4. $A[\varepsilon, \varepsilon]$

Thus $S[t]$ should be derivable if and only if t is the identity in Z^2, and $A[t_1, t_2]$ should be derivable if and only if $t_1 t_2$ is the identity.

The only if part is straightforward. Consider Rule 3 for example. If $x_1 x_2$ and $y_1 y_2$ are the identity, then $t_1 = y_1 x_1 y_2$, $t_2 = x_2$ implies that $t_1 t_2$ is the identity. Etc.

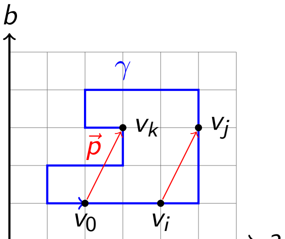
For the if part of Theorem 1 it suffices to show that the grammar above generates $A[u, v]$ for all words u, v such that uv represents the identity in Z^2. The idea is to use Theorem 2 below and argue by induction on $|uv| + \max\{|u|, |v|\}$.

For example consider $(u, v) = (a^3 b^3 a^{-3} b^{-1} a, b^{-1} a^{-2} b^{-1} a)$ with $|uv| + \max\{|u|, |v|\} = 16 + 11 = 27$.

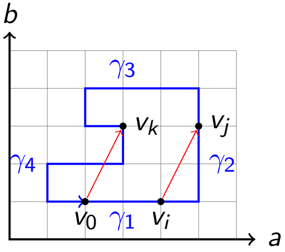


uv is the label of a cycle in the Cayley digram of Z^2 with u being the label from v_0 up to v_k and v the label of the rest of the cycle.

Theorem 2. [Salvati] Consider a simple closed curve γ in Γ with vertices $v_0, v_1, \ldots v_m = v_0$. Let $\vec{p}$ be a vector from v_0 to some interior vertex v_k. If the same vector occurs from v_i to v_j for two internal vertices, then i, j may be chosen so that $0 < i < j < k$ or $k < j < i < m$.



To apply Theorem 2 to the proof of Theorem 1 observe that v_0, v_i, v_j, v_k are the vertices of a parallelogram and that these vertices divide γ into $\gamma_1\gamma_2\gamma_3\gamma_4$ with labels $w_1 = a^2$, $w_2 = ab^2$, $w_3 = ba^{-3}b^{-1}a$, and $w_4 = b^{-1}a^{-2}b^{-1}a$ respectively.



The geometry of the above configuration implies that w_1w_3 and w_2w_4 both define the identity in Z^2. By the induction hypothesis $A[w_1, w_3]$ and $A[w_2, w_4]$ can be generated. An application of Rule 3 generates $A[w_1w_2w_3, w_4] = A[u, v]$.

Conclusion and bibliography

The word problem W for $Z \times Z$ is a 2-MCFL.

Conjecture. Surface groups have multiple context free word problem.

Question. Are groups with MCFL word problem (or well nested MCFL word problem) finitely presented?

Bibliography

Laura Kallmeyer, *Parsing Beyond Context-Free Grammars*, Springer Verlag, 2010.

Sylvain Salvati, MIX is a 2-MCFL and the word problem in Z^2 is captured by the IO and the OI hierarchies, Journal of Computer and System Sciences 81 (2015) 1252–1277.