Efficient Enumeration of Cosets.

Robert Gilman

Algebraic Cryptography Center Web Seminar
Stevens Institute of Technology

November 5, 2009

Coset enumeration

There is no algorithm to decide if a group G given by a finite presentation P is finite.

Coset enumeration

There is no algorithm to decide if a group G given by a finite presentation P is finite.

Coset enumeration gives the correct answer if G is finite.

Coset enumeration

There is no algorithm to decide if a group G given by a finite presentation P is finite.

Coset enumeration gives the correct answer if G is finite.

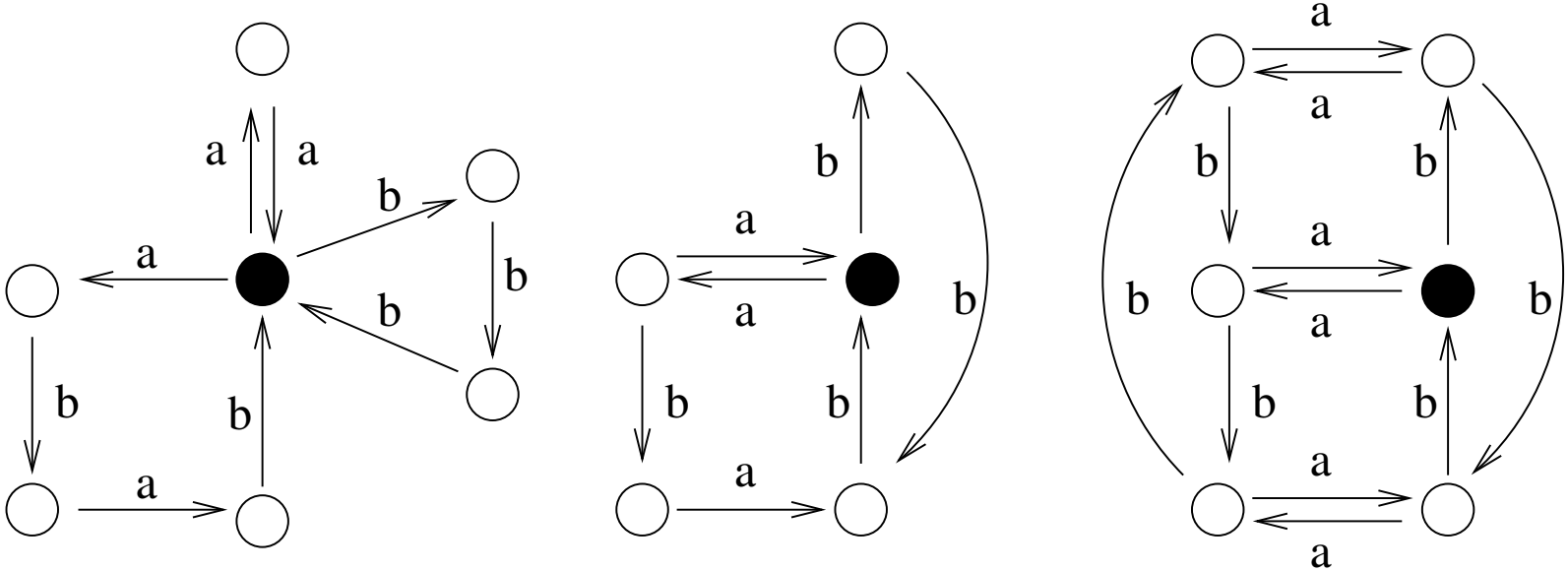
Example: $G = \langle a, b \mid a^2, (ab)^2, b^3 \rangle$

Coset enumeration

There is no algorithm to decide if a group G given by a finite presentation P is finite.

Coset enumeration gives the correct answer if G is finite.

Example: $G = \langle a, b \mid a^2, (ab)^2, b^3 \rangle$



Time complexity

Definition For any presentation P, R is the sum of the lengths of all cyclic conjugates of relators.

The first pass adds at most R new cosets, i.e., vertices.

The second pass adds at most R^2 etc.

After p passes we have at defined at most $O(R^{p+1})$ cosets.

Time complexity

Definition For any presentation P, R is the sum of the lengths of all cyclic conjugates of relators.

The first pass adds at most R new cosets, i.e., vertices.

The second pass adds at most R^2 etc.

After p passes we have at defined at most $O(R^{p+1})$ cosets.

The time complexity is $O(R^{p+1})$ or perhaps $O(R^{3(p+1)})$ taking into account the resolution of coincidences.

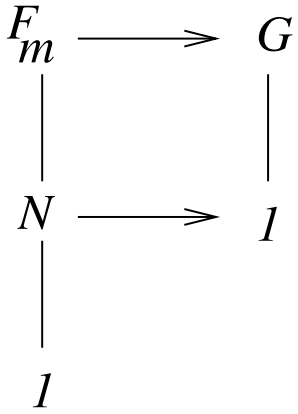
We will measure time by the number of passes.

Estimating the number of passes

Suppose G is finite with presentation $P = \langle a_1, \ldots, a_m \mid r_j \rangle$.

Estimating the number of passes

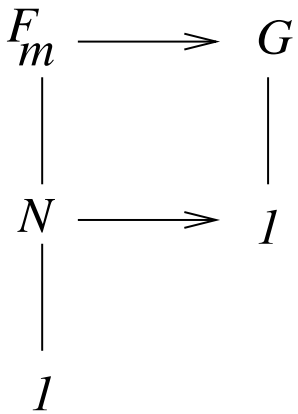
Suppose G is finite with presentation $P = \langle a_1, \ldots, a_m \mid r_j \rangle$.



N is generated by finitely many conjugates of the r_j's.

Estimating the number of passes

Suppose G is finite with presentation $P = \langle a_1, \ldots, a_m \mid r_j \rangle$.



N is generated by finitely many conjugates of the r_j's.

Lemma If N is generated by $\{r_j^w \mid |w| < p\}$, then coset enumeration succeeds in at most $p + 1$ passes.

Worst case complexity

$\{a_1, a_2, \ldots\}$ is a denumerable set of generators.

A presentation has the form $P = \langle a_1, \ldots, a_m \mid r_i \rangle$.

I is the set of all presentations.

J is the subset of I which defines finite groups.

$T(P)$ is the number of passes for $P \in J$.

Worst case complexity

$\{a_1, a_2, \ldots\}$ is a denumerable set of generators.

A presentation has the form $P = \langle a_1, \ldots, a_m \mid r_i \rangle$.

I is the set of all presentations.

J is the subset of I which defines finite groups.

$T(P)$ is the number of passes for $P \in J$.

Let $f(n)$ be an upper bound of $T(P)$ for all $P \in J$ of length at most n.

Worst case complexity

$\{a_1, a_2, \ldots\}$ is a denumerable set of generators.

A presentation has the form $P = \langle a_1, \ldots, a_m \mid r_i \rangle$.

I is the set of all presentations.

J is the subset of I which defines finite groups.

$T(P)$ is the number of passes for $P \in J$.

Let $f(n)$ be an upper bound of $T(P)$ for all $P \in J$ of length at most n.

$f(n)$ is not computable.

Worst case complexity

$\{a_1, a_2, \ldots\}$ is a denumerable set of generators.

A presentation has the form $P = \langle a_1, \ldots, a_m \mid r_i \rangle$.

I is the set of all presentations.

J is the subset of I which defines finite groups.

$T(P)$ is the number of passes for $P \in J$.

Let $f(n)$ be an upper bound of $T(P)$ for all $P \in J$ of length at most n.

$f(n)$ is not computable.

If it were, we could decide whether or not $P \in I$ defined a finite group.

Generic case complexity

Generic case complexity is worst case complexity after discarding a neglible set of unrepresentative difficult instances.

The efficacy of generic case complexity depends on the heuristic observation that difficult instances seem to be rare.

Generic case complexity

Generic case complexity is worst case complexity after discarding a neglible set of unrepresentative difficult instances.

The efficacy of generic case complexity depends on the heuristic observation that difficult instances seem to be rare.

Developers of the Magnus Project observed that recursively unsolvable problems about finitely presented groupws were routinely solved efficiently by relatively simple algorithms.

Generic case complexity

Generic case complexity is worst case complexity after discarding a neglible set of unrepresentative difficult instances.

The efficacy of generic case complexity depends on the heuristic observation that difficult instances seem to be rare.

Developers of the Magnus Project observed that recursively unsolvable problems about finitely presented groupws were routinely solved efficiently by relatively simple algorithms.

I. Kapovich, A. Myasnikov, P. Schupp, V. Shpilrain, Generic-case complexity and decision problems in group theory, J. of Algebra, 264 (2003), 665-694.

I. Kapovich, A. Myasnikov, P. Schupp, V. Shpilrain, Average-case complexity for the word and membership problems in group theory, Advances in Mathematics 190 (2005), 343-359.

The same phenomenon was observed with respect to NP-complete problems in the 1980's and motivated the introduction of average case complexity.

The same phenomenon was observed with respect to NP-complete problems in the 1980's and motivated the introduction of average case complexity.

The Simplex Algorithm has instances which require exponential time, but these instances are never encountered in practice.

Smale and Vershik and Sporyshev showed that the set of exceptional instances has measure 0.

Generic properties of groups

A generic set is the complement of a negligible set.

If G has an infinite cyclic quotient, then the word problem is solvable in linear time on a generic set of inputs. (KMSS)

Generic properties of groups

A generic set is the complement of a negligible set.

If G has an infinite cyclic quotient, then the word problem is solvable in linear time on a generic set of inputs. (KMSS)

If G is finitely presented, then a generic van Kampen diagram has diameter proportional to the log of the length of the word it defines. A generic subset of the set of all words defining the identity can be verified in in polynomial time. (Ushakov)

Generic properties of groups

A generic set is the complement of a negligible set.

If G has an infinite cyclic quotient, then the word problem is solvable in linear time on a generic set of inputs. (KMSS)

If G is finitely presented, then a generic van Kampen diagram has diameter proportional to the log of the length of the word it defines. A generic subset of the set of all words defining the identity can be verified in in polynomial time. (Ushakov)

The generic free basis propety holds if for fixed k a generic set of k-tuples of words in generates a free group (Miasnikov ...

Generic properties of groups

A generic set is the complement of a negligible set.

If G has an infinite cyclic quotient, then the word problem is solvable in linear time on a generic set of inputs. (KMSS)

If G is finitely presented, then a generic van Kampen diagram has diameter proportional to the log of the length of the word it defines. A generic subset of the set of all words defining the identity can be verified in in polynomial time. (Ushakov)

The generic free basis propety holds if for fixed k a generic set of k-tuples of words in generates a free group (Miasnikov ...

This property holds for several classed of groups, but not for Thompson's group (Cleary, Elder, Tabak, ...)

Generic properties of groups

A generic set is the complement of a negligible set.

If G has an infinite cyclic quotient, then the word problem is solvable in linear time on a generic set of inputs. (KMSS)

If G is finitely presented, then a generic van Kampen diagram has diameter proportional to the log of the length of the word it defines. A generic subset of the set of all words defining the identity can be verified in in polynomial time. (Ushakov)

The generic free basis propety holds if for fixed k a generic set of k-tuples of words in generates a free group (Miasnikov ...

This property holds for several classed of groups, but not for Thompson's group (Cleary, Elder, Tabak, ...)

Our approach to coset enumeration follows Ushakov.

Definition of generic sets

J a denumerable set.

$\sigma : J \longrightarrow N_0$, the nonnegative integers is a size function.

Definition of generic sets

J a denumerable set.

$\sigma : J \longrightarrow N_0$, the nonnegative integers is a size function.

$J_0 \subset J_1 \subset \cdots \subset J_n \subset \cdots \subset J$

J_n is the subset of elements of size at most n

Definition of generic sets

J a denumerable set.

$\sigma : J \longrightarrow N_0$, the nonnegative integers is a size function.

$J_0 \subset J_1 \subset \cdots \subset J_n \subset \cdots \subset J$

J_n is the subset of elements of size at most n

μ_n is a probability distribution on J_n

When the J_n's are finite, μ_n is often the equiprobable distribution.

Definition of generic sets

J a denumerable set.

$\sigma : J \longrightarrow N_0$, the nonnegative integers is a size function.

$J_0 \subset J_1 \subset \cdots \subset J_n \subset \cdots \subset J$

J_n is the subset of elements of size at most n

μ_n is a probability distribution on J_n

When the J_n's are finite, μ_n is often the equiprobable distribution.

Definition $X \subset J$ is negligible if $\lim_{n\to\infty} \mu(X \cap J_n) = 0$

Definition of generic sets

J a denumerable set.

$\sigma : J \longrightarrow N_0$, the nonnegative integers is a size function.

$J_0 \subset J_1 \subset \cdots \subset J_n \subset \cdots \subset J$

J_n is the subset of elements of size at most n

μ_n is a probability distribution on J_n

When the J_n's are finite, μ_n is often the equiprobable distribution.

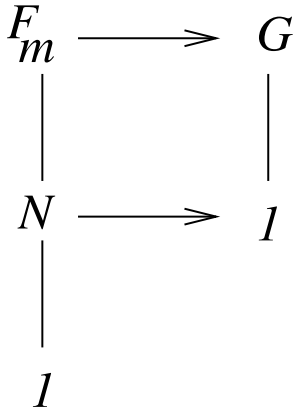
Definition $X \subset J$ is negligible if $\lim_{n\to\infty} \mu(X \cap J_n) = 0$

Definition $X \subset J$ is generic if $\lim_{n\to\infty} \mu(X \cap J_n) = 1$

Generic complexity of coset enumeration

J is the set of finite presentations of finite groups.

We work with a covering $\tilde{J} \longrightarrow J$.



$P \in J$ presents a finite group G

$P = \langle a_1, \ldots, a_m \mid r_j \rangle$

$r_j \in N$

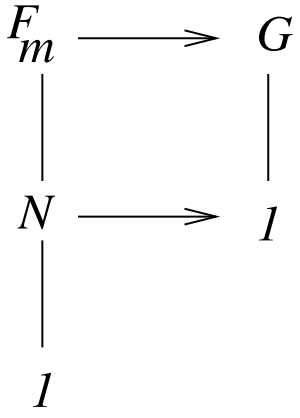
$\{x_k\}$ are Schreier generators for N

x_k is freely equal to $W_k(r_j^{\pm u})$

Generic complexity of coset enumeration

J is the set of finite presentations of finite groups.

We work with a covering $\tilde{J} \longrightarrow J$.



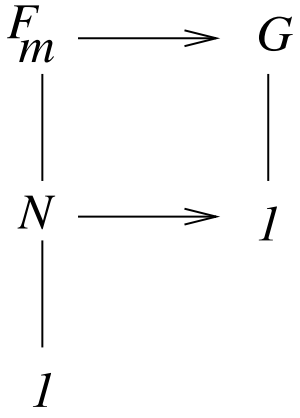
$P \in J$ presents a finite group G
$P = \langle a_1, \ldots, a_m \mid r_j \rangle$
$r_j \in N$
$\{x_k\}$ are Schreier generators for N
x_k is freely equal to $W_k(r_j^{\pm u})$

$\tilde{J} = \{(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$, $r_j \in N$, $u_t \in F_m$

Generic complexity of coset enumeration

J is the set of finite presentations of finite groups.

We work with a covering $\tilde{J} \longrightarrow J$.



$P \in J$ presents a finite group G
$P = \langle a_1, \ldots, a_m \mid r_j \rangle$
$r_j \in N$
$\{x_k\}$ are Schreier generators for N
x_k is freely equal to $W_k(r_j^{\pm u})$

$$\tilde{J} = \{(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\},\ r_j \in N,\ u_t \in F_m$$

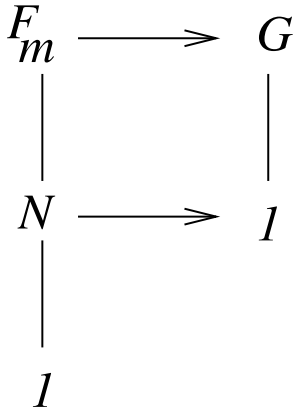
$$(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\}) \longrightarrow \langle a_i \mid r_j, x_k(W_k(r_p^{\pm u_p})^{-1} \rangle$$

with all relators cyclicly reduced.

Generic complexity of coset enumeration

J is the set of finite presentations of finite groups.

We work with a covering $\tilde{J} \longrightarrow J$.



$P \in J$ presents a finite group G
$P = \langle a_1, \ldots, a_m \mid r_j \rangle$
$r_j \in N$
$\{x_k\}$ are Schreier generators for N
x_k is freely equal to $W_k(r_j^{\pm u})$

$\tilde{J} = \{(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$, $r_j \in N$, $u_t \in F_m$

$$(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\}) \longrightarrow \langle a_i \mid r_j, x_k(W_k(r_p^{\pm u_p})^{-1} \rangle$$

with all relators cyclicly reduced.

Lemma Every $P \in J$ is obtained in this way.

$$(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\}) \to \langle a_i \mid r_j, x_k(W_k(r_p^{\pm u_p})^{-1} \rangle$$

Lemma $\tilde{J}$ maps onto J, and P presents F_m/N.

$$(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\}) \to \langle a_i \mid r_j, x_k(W_k(r_p^{\pm u_p})^{-1} \rangle$$

Lemma $\tilde{J}$ maps onto J, and P presents F_m/N.

Proof Coset enumeration. Every $r_p^{u_p}$ can eventually be traced at 1.

Hence each $W_k(r_p^{u_p}, \ldots)$ can be traced.

Since $x_k W_k(r_p^{u_p, \cdots})^{-1}$ is a relator, x_k is traceable.

All relators of P lie in N and every freely reduced word in N is traceable. $\square$

The number of passes is the maximum of the lengths of the u_p's plus 1.

The size function

$$(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\}) \to \langle a_i \mid r_j, x_k(W_k(r_p^{\pm u_p})^{-1} \rangle$$

The size of an element of $\tilde{J}$ is the maximum of

1. The number of words r_j
2. The number of words u_t
3. The lengths of the words W_k, r_j, u_t

The size function

$$(m, N, \{x_k\}, \{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\}) \to \langle a_i \mid r_j, x_k(W_k(r_p^{\pm u_p})^{-1} \rangle$$

The size of an element of $\tilde{J}$ is the maximum of

1. The number of words r_j
2. The number of words u_t
3. The lengths of the words W_k, r_j, u_t

The length of the r_j's is as words in the Schreier generators of N.

When they are rewritten as words in F_m, they do not get any shorter.

J_n is the image of $\tilde{J}_n$

The probability distributions

The probability distribution μ_n is induced by $\tilde{\mu}_n$ on $\tilde{J}_n$.

$\tilde{J}_n = \cup_{m,N}(m, N, \{x_k\}) \times \{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

The probability distributions

The probability distribution μ_n is induced by $\tilde{\mu}_n$ on $\tilde{J}_n$.

$\tilde{J}_n = \cup_{m,N}(m, N, \{x_k\}) \times \{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

Use any distribution ν on $\cup_{m,N}(m, N, \{x_k\})$

Use the equiprobable distribution on $\{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

The probability distributions

The probability distribution μ_n is induced by $\tilde{\mu}_n$ on $\tilde{J}_n$.

$\tilde{J}_n = \cup_{m,N}(m, N, \{x_k\}) \times \{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

Use any distribution ν on $\cup_{m,N}(m, N, \{x_k\})$

Use the equiprobable distribution on $\{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

The r_j's are chosen uniformly at random from words on length at most n in the Schreier generators of N

The probability that all r_j's have length less than $.9n$ goes 0 very fast

The probability distributions

The probability distribution μ_n is induced by $\tilde{\mu}_n$ on $\tilde{J}_n$.

$\tilde{J}_n = \cup_{m,N}(m, N, \{x_k\}) \times \{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

Use any distribution ν on $\cup_{m,N}(m, N, \{x_k\})$

Use the equiprobable distribution on $\{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

The r_j's are chosen uniformly at random from words on length at most n in the Schreier generators of N

The probability that all r_j's have length less than $.9n$ goes 0 very fast

The set of elements in $\tilde{J}$ such that the maximum length of the relators is at least .9 of the maximum of the lengths of the u_t's is generic

The probability distributions

The probability distribution μ_n is induced by $\tilde{\mu}_n$ on $\tilde{J}_n$.

$\tilde{J}_n = \cup_{m,N}(m, N, \{x_k\}) \times \{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

Use any distribution ν on $\cup_{m,N}(m, N, \{x_k\})$

Use the equiprobable distribution on $\{(\{W_k(x_p, y_p)\}, \{r_j\}, \{u_t\})\}$

The r_j's are chosen uniformly at random from words on length at most n in the Schreier generators of N

The probability that all r_j's have length less than $.9n$ goes 0 very fast

The set of elements in $\tilde{J}$ such that the maximum length of the relators is at least .9 of the maximum of the lengths of the u_t's is generic

Theorem Let L be the maximum relator length. For presentations in J, coset enumeration succeeds in at most $1.2L$ passes on a generic set of presentations.

Discussion

1. The probability distributions μ_n are chosen in an ad hoc way to make the argument work. How useful is the result?

Discussion

1. The probability distributions μ_n are chosen in an ad hoc way to make the argument work. How useful is the result?
2. Charlie Sims observed that the presentation $\langle a_1, \ldots, a_m \mid (a_i a_j)^2 \rangle$ with $L = 2$ requires more than 3 passes for all but finitely many m, so these presentations are not included in the generic set. In fact the set of all presentations with relators of at most some contstant length is negligible.

Discussion

1. The probability distributions μ_n are chosen in an ad hoc way to make the argument work. How useful is the result?
2. Charlie Sims observed that the presentation $\langle a_1, \ldots, a_m \mid (a_i a_j)^2 \rangle$ with $L = 2$ requires more than 3 passes for all but finitely many m, so these presentations are not included in the generic set. In fact the set of all presentations with relators of at most some contstant length is negligible.
3. Where are the difficult presentations? Perhaps we should look at presentations with short relators.

Discussion

1. The probability distributions μ_n are chosen in an ad hoc way to make the argument work. How useful is the result?
2. Charlie Sims observed that the presentation $\langle a_1, \ldots, a_m \mid (a_i a_j)^2 \rangle$ with $L = 2$ requires more than 3 passes for all but finitely many m, so these presentations are not included in the generic set. In fact the set of all presentations with relators of at most some contstant length is negligible.
3. Where are the difficult presentations? Perhaps we should look at presentations with short relators.
4. On the other hand B. H. Neumann found a series of presentations P_n of the trivial group with 3 generators, 3 relators and $L = 5^n$. The generation of this sequence seems somewhat similar to the product replacement procedure for generating random group elements. Are these P_n's random presentations? Does coset enumeration converge in $1 + 5^n$ passes on P_n?