# Automatic Groups and String Rewriting

Robert H. Gilman

## 1 Introduction

In recent years developments in geometric topology have led to a remarkable interplay between group theory, geometry, and the theory of automata and formal languages. One of the best known result of this interplay has been the introduction of automatic groups. In this article we will consider this class of groups from the point of view of automata and language theory. We will see how work of Epstein, Holt and Rees [6] puts automatic groups naturally into the context of string rewriting. We assume that the reader is familiar with the Knuth-Bendix completion procedure for string rewriting [10], [12], [19, Chapter 2]. For a guide to the topological developments referred to above consult [5, Chapter 12] and [18].

Let $G$ be a finitely generated group. In what follows we will distinguish between words in the generators and elements of the group. A *choice of generators* for $G$ is a map $\pi : \Sigma \to G$ which extends to a surjective monoid homomorphism $\pi : \Sigma^* \to G$. Here $\Sigma$ is a finite alphabet and $\Sigma^*$ is the free monoid on $\Sigma$. We require that $\Sigma$ be equipped with formal inverses and that $\pi$ respect these inverses. More precisely $\Sigma = \{a, a^{-1}, b, b^{-1} \ldots\}$, and the involutory permutation $a \leftrightarrow a^{-1}, b \leftrightarrow b^{-1}, \cdots$ extends uniquely to an anti-automorphism $w \to w^{-1}$ of $\Sigma^*$. The homomorphism $\pi$ satisfies $\pi(w^{-1}) = (\pi(w))^{-1}$.

## 2 String Rewriting with Infinite Sets of Rules

Suppose $G = \langle a, b, \ldots \mid w_1 = v_1, w_2 = v_2, \ldots \rangle$ is a finitely presented group, $\Sigma = \{a, a^{-1}, b, b^{-1}, \ldots\}$, and $<$ is a reduction ordering on $\Sigma^*$ suitable for the Knuth-Bendix procedure. That is, $<$ is a well-ordering which respects multiplication.

**Example 1** $\Sigma = \{a, a^{-1}, b, b^{-1}\}$*,* $G = \langle a, b \mid ab = ba \rangle$*, and* $<$ *is the length plus lexicographic order with* $a < b < b^{-1} < a^{-1}$*.*

Clearly the group $G$ of Example 1 is isomorphic to $Z \times Z$, the direct product of the integers with itself. The KB procedure must fail to terminate in this example because there is no finite complete rewriting system for $G$ with the given generators and reduction ordering. An infinite complete system is given in Table 1. It is easy to check these assertions. The infinite system is complete

| | |
|---|---|
| $aa^{-1} \to \epsilon$ | $a^{-1}a \to \epsilon$ |
| $bb^{-1} \to \epsilon$ | $b^{-1}b \to \epsilon$ |
| $ba \to ab$ | $b^{-1}a \to ab^{-1}$ |
| $a^{-1}b \to ba^{-1}$ | $a^{-1}b^{-1} \to b^{-1}a^{-1}$ |
| $ab^na^{-1} \to b^n, n \geq 1$ | $ab^{-n}a^{-1} \to b^{-n}, n \geq 1$ |

Table 1: An infinite complete rewriting system for Example 1.

because the language of irreducible words, $\{a^ib^j, b^ja^{-i} \mid i, j \in Z, i \geq 0\}$, is a set of normal forms for elements of $G$. If there were a finite complete rewriting system, then each of its reductions would be a consequence of a finite number of the reductions in Table 1. Hence some finite subset of the reductions in this table would be a complete rewriting system. But the system in Table 1 is reduced in the sense that the left hand side of every rewrite rule is irreducible with respect to all the other rules. It follows that any finite subsystem would have more irreducible words than the whole system. Thus finite subsystems cannot be complete.

In practice the KB procedure may generate a sequence of rewrite rules which looks like the start of an infinite system similar to the one in Table 1, and it is natural to try to extend the procedure to deal with this situation. Let us identify rewrite rules $w \to v$ with ordered pairs $(w, v) \in \Sigma^* \times \Sigma^*$. We might try to extend the KB procedure to systems of rewrite rules which are rational subsets of $\Sigma^* \times \Sigma^*$, but completeness is undecidable for these systems even if they are known to be terminating [1, Chapter III.1.3], [16]. The method we will use makes use of the smaller class of synchronized rational sets. As we mentioned above, it is taken from [6]. Other methods are discussed in [1] and a method using term rewriting has been implemented in [15].

**Definition 1** *A synchronized rational relation (or synchronized relation for short) on* $\Sigma^* \times \Sigma^*$ *is one accepted by a two-tape automaton in which both heads move synchronously. Equivalently it is a regular subset of* $(\Sigma \times \Sigma)^* \big((\Sigma^* \times \{\epsilon\}) \cup (\{\epsilon\} \times \Sigma^*)\big)$.

The infinite rewriting system of Table 1 is a synchronized relation. For example $\{ab^na^{-1} \to b^n \mid n \geq 1\} = (a, b)(b, b)^*(b, \epsilon)(a, \epsilon)$. See [5, Chapter 1] and [7] for a review of the properties of synchronized relations. Boolean operations on synchronized relations are computable in theory and in practice, and so are equality and composition. In addition $n$-ary synchronized relations are closed under existential and universal quantification, and the corresponding theory is decidable.

We may imagine that we run the KB procedure on Example 1 for a finite amount of time and obtain the rules in Table 2. We will use this example to illustrate our extension of the KB procedure. The first step is to infer from a

| | |
|---|---|
| $aa^{-1} \to \epsilon$ | $a^{-1}a \to \epsilon$ |
| $bb^{-1} \to \epsilon$ | $b^{-1}b \to \epsilon$ |
| $ba \to ab$ | $b^{-1}a \to ab^{-1}$ |
| $a^{-1}b \to ba^{-1}$ | $a^{-1}b^{-1} \to b^{-1}a^{-1}$ |
| $aba^{-1} \to b$ | $ab^{-1}a^{-1} \to b^{-1}$ |
| $ab^2a^{-1} \to b^2$ | $ab^{-2}a^{-1} \to b^{-2}$ |

Table 2: A finite incomplete rewriting system for Example 1.

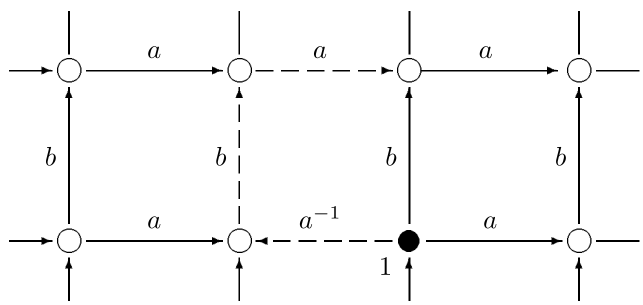


Figure 1: The path $a^{-1}ba$ in the Cayley diagram for Example 1.

finite set of rewrite rules like those in Table 2 a valid infinite set in the form of a synchronous relation, and the second step is to verify that this infinite set is complete.

## 2.1 A Rule of Inference

To accomplish the first step mentioned above we make use of a construction based on the Cayley diagram $\Gamma$ of $G$. First we note that words in $\Sigma$ may be viewed as paths starting at the identity in $\Gamma$ as in Figure 1. We will identify words with their corresponding paths. For any word $w \in \Sigma^*$ and integer $i \geq 0$, define $w(i)$ to be the prefix of $w$ length $\min\{i, |w|\}$ where $|w|$ denotes the length of $w$.

**Definition 2** *For $w, v \in \Sigma^*$ the word differences from $w$ to $v$ are the group elements $g_i = \pi(w(i)^{-1}v(i))$.*

Figure 2 suggests a path consisting of a sequence of group elements $1, g_1, g_2, g_3, g_4$ joined by edges with labels $(a, b), (b, b), (b\epsilon), (a^{-1}, \epsilon)$.

**Definition 3** *The Cayley automaton $\Gamma_{g'}$ corresponding to a choice of generators $\pi : \Sigma \to G$ and a choice of $g' \in G$ is the automaton with vertices $g \in G$ and*

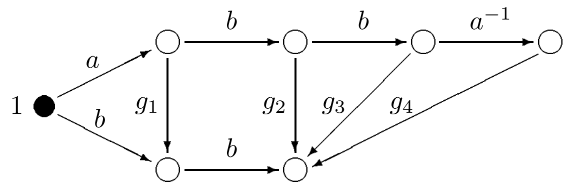


Figure 2: Word differences from $abba^{-1}$ to $bb$

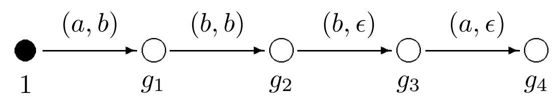


Figure 3: The path in the Cayley automaton corresponding to the word differences in Figure 2.

*an edge from $g$ to $h$ with label $(a, b)$ whenever $\pi(a^{-1})g\pi(b) = h$ for $a, b \in \Sigma \cup \{\epsilon\}$. The initial vertex is 1 and the single terminal vertex is $g'$.*

The vertices are the states of the automaton and the labeled edges indicate the transitions on a given input. In general $\Gamma_{g'}$ is infinite. It is straightforward to show that it accepts $\{(w, v) \mid \pi(w^{-1}v) = g'\}$.

Assume from now on that the reduction ordering $<$ is a length and lexicographic ordering; this is the case considered in [6]. Given any finite set of valid rewrite rules, we can infer a possibly infinite set by constructing a synchronized automaton which has part of $\Gamma_1$ as a homomorphic image. Usually we do not know $G$ directly and so cannot construct $\Gamma_1$ itself. Instead for each rewrite rule $w \to v$ we reduce the words $w(i)^{-1}v(i)$ to irreducible words $u(i)$ and use the $u(i)$'s as vertices in place of the word differences $\pi(w(i)^{-1}v(i))$. In order to insure that our automaton is synchronous we will have two sets of vertices, $V_0$ and $V_1$. As the same word $w$ may appear in both sets, we will denote its appearance in the second set by $\tilde{w}$ to avoid confusion. Edges with labels in $\Sigma \times \Sigma$ will connect pairs of vertices in $V_0$, and edges with labels in $\Sigma \times \{\epsilon\}$ will have their target vertices in $V_1$. There will be no other edges.

**Definition 4** *Let $\pi : \Sigma \to G$ be a choice of generators, $<$ a length and lexicographic reduction ordering on $\Sigma^*$, and $\mathcal{R}$ a finite reduced set of rewrite rules with respect to $<$. Suppose $a \in \Sigma \cup \{\epsilon\}$. The synchronous automaton $A_a$ is constructed as follows. For each rule $w \to v \in \mathcal{R}$ with $w = a_1 \cdots a_n$ and $v = b_1 \cdots b_m$, reduce the words $w(i)^{-1}v(i)$ to irreducible words $u(i)$. Put $u(i)$ in*

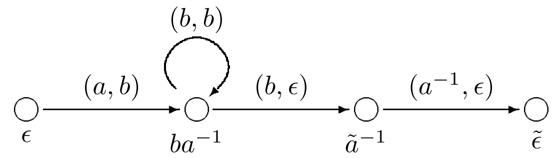


Figure 4: Part of the automaton $A_\epsilon$ for Example 1. The edges shown are determined by the rewrite rule $ab^2a^{-1} \to b^2$ from Table 2.

*$V_0$ if $i \le m$ and in $V_1$ if $i > m$. Add an edge*

$$\begin{cases} u(i-1) \overset{(a_i,b_i)}{\to} u(i) & \text{if } 0 < i \le m \\ u(i-1) \overset{(a_i,\epsilon)}{\to} \tilde{u}(i) & \text{if } i = m+1 \\ \tilde{u}(i-1) \overset{(a_i,\epsilon)}{\to} \tilde{u}(i) & \text{if } m+1 < i \le n. \end{cases}$$

*The initial vertex of $A_a$ is $\epsilon \in V_0$ and the terminal ones are $a \in V_0$ and $a \in V_1$.*

Because of the choice of $<$ we know $m \le n$ in the definition above. It is clear that $A_a$ accepts a synchronized relation. Figure 4 shows the part of $A_\epsilon$ from Example 1 which accepts the infinite set of rewrite rules $ab^na^{-1} \to b^n, n \ge 1$.

In practice one runs the KB procedure until the set of word differences stabilizes and then constructs the automata $A_a$. The proof of the following theorem is straightforward from the construction of the Cayley automaton.

**Theorem 1** *There is a homomorphism from the automaton $A_a$ to the Cayley automaton $\Gamma_{\pi(a)}$.*

**Corollary 1** *If $A_a$ accepts $(w, v)$, then $\pi(wa) = \pi(v)$. In particular the rewrite rules accepted by the automaton $A_\epsilon$ are valid.*

## 2.2 Verifying confluence

Suppose $\mathcal{R}^+$ is the possibly infinite synchronous set of rewrite rules accepted by an automaton $A_\epsilon$ constructed as above. How do we verify that $\mathcal{R}^+$ is complete? Let $\rho_a$ be the synchronous relation accepted by $A_a$. If $L$ is the language of words irreducible with respect to $\mathcal{R}^+$, then $\mathcal{R}^+$ will be complete if and only if $L$ projects bijectively to $G$. Since $L$ is rational, we can use the decidability of synchronous relations to check that the following conditions hold.

1. The composition $\rho_a \circ \rho_{a^{-1}}$ is the identity on $L$;

2. For each relation $a_1 \dots a_n = b_1 \dots b_m$ in the finite presentation of $G$, $\rho_{a_1} \circ \dots \circ \rho_{a_n} = \rho_{b_1} \circ \dots \circ \rho_{b_m}$ on $L$.

If these conditions hold, then each $\rho_a$ is a permutation of $L$, and the group generated by these permutations is isomorphic to $G$. It follows that $L$ projects bijectively to $G$ and the rewrite rules accepted by $A_\epsilon$ are complete. See [6] for the details. If the conditions do not hold, we can run the KB procedure longer and try again.

In this context it is natural to consider the class of groups $G$ with generators $\pi : \Sigma \to G$ and a regular set $L \subset \Sigma$ such that $L$ projects bijectively to $G$ and for all $a \in \Sigma$ the relation $w \sim_a v$ if $w, v \in L$ and $\pi(wa) = \pi(v)$ is synchronized. This is the essentially the class of an automatic groups.

**Definition 5** *Let $G$ be a finitely generated group with choice of generators $\pi : \Sigma \to G$. $G$ is automatic if there is a regular set $L \subset \Sigma^*$ such that $\pi(L) = G$ and if for each $a \in \Sigma \cup \{\epsilon\}$ the binary relation $\sim_a$ defined by $w \sim_a v$ if $w, v \in L$ and $\pi(wa) = \pi(v)$ is a synchronized rational relation.*

We required above that $L$ project bijectively instead of merely surjectively as Definition 5 allows; but by [5, Theorem 2.5.1] if $G$ is automatic, then we may take $L$ to project bijectively (in which case $\sim_\epsilon$ is just the identity on $L$). We shall say that $L$ is part of an automatic structure for $G$ if the conditions of Definition 5 hold.

# 3 The geometric view of automatic groups.

Suppose $G$ is finitely generated with choice of generators $\pi : \Sigma \to G$ and $\Gamma$ is the corresponding Cayley diagram. There is a well known metric $d$ on $\Gamma$ which assigns to each pair of points the distance of the shortest path between them. With this definition we can define the distance between $w$ and $v$ to be

$$D(w, v) = \max_i d\big(\pi(w(i)), \pi(v(i))\big).$$

$D(w, v)$ is essentially the maximum separation achieved by two particles starting from 1 at the same time and moving along the paths determined by $w$ and $v$ at unit speed. More precisely it is the maximum separation achieved at the vertices along the path. $D$ is not a metric since it is possible for different words to determine the same path, but $D$ is symmetric and satisfies the triangle inequality. One of the first results in the theory of automatic groups is the following geometric characterization [5, Theorem 2.3.5].

**Theorem 2** *Suppose that $G$ is a finitely generated group with choice of generators $\pi : \Sigma \to G$. $G$ is automatic if and only if there is a language $L \subset \Sigma^*$ and constant $K$ such that*

1. *$L$ is rational;*
2. *$\pi(L) = G$;*

*3. For any $w, v \in L$ with $d(\pi(w), \pi(v)) \leq 1$, $D(w, v) \leq K$.*

Condition 3 is sometimes called the $K$ fellow traveller property. If $G$ satisfies the conditions of Theorem 2 with respect to one set of generators, then it does so with respect to any other set of generators although the value of $K$ may change. There are a number of interesting language theoretic problems arising from the theory of automatic groups. Perhaps the most intriguing one is that it is not known if condition 1 has any effect. That is, if we omit condition 1 in Definition 2, do we get any more groups? Let us call a group $G$ *combable* if there exists a set $L \subset \Sigma^*$ for which $G$ satisfies conditions 2 and 3. (The definition of combable in [5] is more restrictive.) It is known that $H_3$, the group of all integer matrices of the form

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

is not automatic, but it is not known whether or not it is combable.

There is another measure of distance between paths, $D'(w, v)$, which is the maximum separation (at vertices) achieved by two particles starting from 1 at the same time and moving along the paths determined by $w$ and $v$ with each particle moving at variable but nonnegative speed in such a way as to minimize the maximum separation. A finitely generated group $G$ which satisfies the conditions of Theorem 2 with $D'$ in place of $D$ is called *asynchronously automatic* [5, Chapter 7]. In the case of asynchronously automatic groups it is known that relaxing the condition that $L$ be regular allows more groups. Specifically $H_3$ not asynchronously automatic [5, Theorem 8.2.8] and is not included if we allow $L$ to be context free, but it is included if we allow $L$ to range over the larger class of indexed languages [2].

# 4 Hyperbolic Groups

Hyperbolic group are an interesting class of groups introduced by Gromov [11]. Other references are [9] and [4].

**Definition 6** *Let $G$ be finitely generated with choice of generators $\pi : \Sigma \to G$ and Cayley diagram $\Gamma$. A* geodesic *is a shortest path between two points in $\Gamma$, and a* geodesic triangle *is a triangle whose sides are geodesics. $G$ is* hyperbolic *if there is a constant $\delta$ such that for every geodesic triangle each point on any side is a distance at most $\delta$ from some point on one of the other two sides.*

The validity of the triangle condition above is independent of the choice of generators although the value of $\delta$ is not. All groups satisfying the small cancellation hypothesis $C'(1/6)$ or the hypotheses $C'(1/4)$ and $T(4)$ are hyperbolic [20] but it is easy to check that $Z \times Z$ is not.

Hyperbolic groups are a proper subclass of automatic groups. If $\pi : \Sigma \to G$ is any choice of generators for the hyperbolic group $G$, then the language $L$ of all words $w \in \Sigma^*$ whose corresponding paths are geodesics is a regular set [11, §5] and is part of an automatic structure for $G$ as in Definition 5 [5, Theorem 3.4.5]. In fact this condition characterizes hyperbolic groups. That is, $G$ is hyperbolic if and only if its language of geodesics is rational and is part of an automatic structure for $G$ [17].

Now we make a connection with string rewriting. A survey of groups with various kinds of length reducing confluent presentations is given in [13], and hyperbolic groups fill one gap in the table given there. They are precisely the groups which admit a length reducing presentation which is confluent on the identity [4, Chapter 4, §5]. It follows that a group which admits a length reducing rewriting system confluent at the identity with respect to one choice of generators does so with respect to any choice of generators. It does not seem easy to prove this result directly.

From the point of view of language theory a group is hyperbolic if and only if the language of words representing the identity (sometimes called the word problem of $G$) is generated by a context sensitive grammar in which productions are strictly length increasing and in which every nonterminal derives a terminal. These languages are included in the abstract family of growing context sensitive languages [3]. This AFL includes context free languages, and groups whose word problems are context free are known to be the groups with a free subgroup of finite index [14]. It would be interesting to characterize groups whose word problems are growing context sensitive.

# References


[1] B. Benninghofen, S. Kemmerich, M. M. Richter, *Systems of Reductions*, Lecture Notes in Computer Science **277**, Springer Verlag, Berlin 1987.

[2] M. Bridson and R. Gilman, Formal language theory and the geometry of 3-manifolds, to appear.

[3] G. Buntrock and K. Lorys, On growing context-sensitive languages, 19th ICALP, Lecture Notes in Computer Science **623**, Springer Verlag, Berlin 1992, 77-88.

[4] M. Coornaert, T. Delzant, and A. Papadopoulos, *Géométrie et théorie des groupes*, Lecture Notes in Mathematics **1441**, Springer Verlag, Berlin 1990.

[5] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston, *Word Processing in Groups*, Jones and Bartlett, Boston 1992.

[6] D. B. A. Epstein, D. F. Holt, and S. Rees, The use of Knuth-Bendix methods to solve the word problem in automatic groups, J. Symbolic Computation **12** 1991, 397–414.

[7] C. Frougny and J. Sakarovitch, Synchronized rational relations of finite and infinite words, Theoretical Computer Science **108** 1993, 45–82.

[8] S. M. Gersten and H. B. Short, Rational subgroups of biautomatic groups, Annals of Mathematics **134** 1991, 125–158.

[9] E. Ghys and P. de la Harpe eds., *Sur les Groupes Hyperboliques d'après Mikhael Gromov*, Birkhäuser, Boston 1990.

[10] R. Gilman, Presentations of groups and monoids, J. Algebra **57** 1979, 544–554.

[11] M. Gromov, Hyperbolic groups, in *Essays in Group Theory*, S. M. Gersten ed., MSRI Publ. **8**, Springer Verlag, New York 1987, 75–263.

[12] M. Jantzen, *Confluent String Rewriting*, EATCS Monographs on Theoretical Computer Science **14**, Springer Verlag, Berlin 1988.

[13] K. Madlener and F. Otto, Groups presented by certain classes of finite length-reducing string rewriting systems, Lecture Notes in Computer Science **256**, Springer Verlag, Berlin 1987, 133–144.

[14] D. Muller and P. Schupp, Groups, the theory of ends and context-free languages, J. Computer and System Sciences **26** 1983, 295–310.

[15] R. Needham, Term rewriting and the word problem for certain infinite presentations of groups, Dissertation, University of Michigan 1992.

[16] C. O'Dunlaing, Infinite regular Thue systems, Theor. Comp. Sci. **25** 1983, 171–192.

[17] P. Papasoglu, Geometric methods in group theory, Dissertation, Columbia Univ., 1993

[18] G. P. Scott, The geometries of 3-manifolds, Bulletin of the London Mathematical Society, **15** 1983, 401–487.

[19] C. Sims, *Computation with Finitely Presented Groups* Encyclopedia of Math. and Its Applications **48**, Cambridge Univ. Pr. 1994.

[20] R. Strebel, Small cancellation groups, in *Sur les Groupes Hyperboliques d'après Mikhael Gromov*, E. Ghys and P. de la Harpe eds., Birkhäuser, Boston 1990, 227–273.