Automatic Groups: A Guided Tour

Benson Farb *

appeared in *L'Enseignement Math.*, Vol. 38 (1992), pp. 291-313.

Geometric group theory has seen a flurry of activity in the last ten years due to new geometric ideas of Gromov, Cannon, Thurston and others. This resurgence has given birth to a truly new field of mathematics: the theory of automatic groups. This fast-growing field lies at the juncture of geometry, topology, combinatorial group theory and algorithms; many of its ideas and themes have their roots in mathematics from throughout this century. As this subject is nearing the end of its infancy it seems an appropriate time for a quick survey. The writing of this paper grew out of conversations with the curious, talks given at Cornell and Princeton, and an attempt to set things straight in my own mind. I'll try to give the reader a taste of some of the main ideas, techniques and applications of the theory. For a detailed, comprehensive introduction to automatic groups, the reader may consult the upcoming book *Word Processing and Group Theory* by Epstein, Cannon, Holt, Levy, Paterson and Thurston ([E et al.]).

1 Background in Geometric Group Theory

The theory of automatic groups is based on the study of groups from a geometric viewpoint. In order to put things in their proper perspective we'll first have to review a small bit of background material. To a group G with finite generating set S, one associates the *Cayley graph* $\Gamma_S(G)$ of (G,S), which is a directed graph whose vertex set consists of elements of G, with a directed edge labelled s going from g to $g \cdot s$ for each $g \in G, s \in S$. As a matter of convenience, for elements $s \in S$ which have order two, we draw only one (undirected) edge labelled s between g and $g \cdot s$, as opposed to drawing one from g to $g \cdot s$ and another from $g \cdot s$ to g. The Cayley

*Supported by a National Science Foundation Graduate Fellowship

Figure 1: Cayley Graphs of Z, $Z \times Z$ and $Z * Z$ with the standard generators.

graph can be made into a metric space by assigning each edge length 1, and by defining the distance between two points to be the length of the shortest path between them. A word representing $g \in G$ can be thought of as a path in $\Gamma_S(G)$ from 1 to g; a geodesic in $\Gamma_S(G)$ from g to h is the same as a word of minimal length representing $g^{-1}h$. Examples of Cayley graphs of some familiar groups are given in figure 1. It is possible to construct and study geometric objects such as geodesics and triangles in the metric space $\Gamma_S(G)$. One theme of geometric group theory is that there is an interplay between geometric properties of the Cayley graph and group theoretic properties of G.

For general finitely presented groups (i.e., groups having a presentation with finitely many generators and finitely many relators), not much can be said. Recall that if G is a finitely presented group with generating set S, the "word problem" for G consists of giving an algorithm which takes as input any two elements of the free group on S (so-called 'words'), and as output tells whether or not the two words represent the same element of G; equivalently, there is an algorithm which tells whether or not an input word represents the identity element of G. In one of the great mathematical achievements of the 1950's, Novikov and Boone found a finitely presented group for which the word problem is not solvable. We shouldn't give up so easily, though, since we are mostly interested in studying groups that arise in geometry and topology, and it is in these situations where we can hope to use the structure of the spaces to tell us more about the groups.

If G arises naturally from some geometric situation, for example if G is the fundamental group of some compact hyperbolic manifold M, then the geometry of the Cayley graph $\Gamma_S(G)$ is in some sense a combinatorial model of the geometry of M. Geodesics, spaces at infinity, and global manifestations of curvature are all captured by the metric space $\Gamma_S(M)$. In fact, if G is the fundamental group of a compact Riemannian manifold M, then $\Gamma_S(G)$ is quasi-isometric (i.e., isometric up to constant factors) to the universal cover $\widetilde{M}$ of M. The study of the geometry of the Cayley graph and its group theoretic implications is part of the field of geometric group theory. For an inspiring introduction to some of this material, see [Ca2]. As Cannon notes in his paper, one of the central philosophies of geometric group theory is that geometric models of groups give rise to computational schemes for dealing with those groups. It is this idea that we shall explore.

2 The Beginnings of Automatic Groups

The main ideas behind automatic groups began with another beautiful paper of Cannon. In the second decade of this century Dehn solved the word problem for surface groups by using the geometry of the hyperbolic plane. In 1984 Cannon extended Dehn's solution to the word problem to cocompact discrete groups of hyperbolic isometries ([Ca1]). Perhaps the main idea of Cannon's paper is that one can "see" the Cayley graphs of such groups. What does it mean to see the Cayley graph of an infinite group? For example, the Cayley graph of (Z, S) with $S = \{1\}$ is simply the real line with a vertex at each integer. When drawing a picture of $\Gamma_S(Z)$, we draw only the two-ball, say, and then a trailing line of dots $\cdots$. By this we really mean to repeat the picture of the two-ball in our heads off to infinity, thinking also of the linear recursion $n \longmapsto n+1$. To "see" the Cayley graph should mean to have a picture of some finite ball around the origin and some finite machine which tells us how to piece copies of this ball together out to infinity. Cannon suggested as an open problem that one might "Formalize the notion that a Cayley graph can be described by linear recursion, and devise efficient algorithms for working out that recursion for many examples." The idea is that if such a linear recursion exists, which should happen whenever there is some pattern in the Cayley graph, then we can build a picture of what the group looks like, and from this picture we can construct algorithms to do computations in the group, such as solving the word problem.

The next layer of foundation was provided by Thurston, who gave a formal definition of the "linear recursion" Cannon spoke of. Thurston did this by using finite state automata (FSA for short), the simplest type of machines which have been studied thoroughly by computer scientists for nearly forty years. It seems interesting that Gilman was independently exploring the use of FSA for normal forms in groups (see, e.g. [Gi]), although with no (explicit) discussion of geometry. The details of the basic theory of automatic groups were worked out at Warwick by Epstein, Holt and Paterson. The use of finite state automata is partly motivated by their success in both the theory and applications of computer science; most word-processors (including 'vi') construct finite state automata for tasks such as word searches, and many compilers use FSA during lexical and syntactical analysis. In order to understand automatic groups we'll first need to have some understanding of finite state automata.

3 Finite State Automata

Given some finite set of letters $\mathcal{A} = \{a_1, \ldots, a_n\}$, we want to pick out a nice subset of the set $\mathcal{A}^*$ of all words in the letters a_i (one can view $\mathcal{A}^*$ as the free monoid generated by the elements of $\mathcal{A}$). A subset $L \subseteq \mathcal{A}^*$ is called a *language*. Informally, a *finite state automaton* W over $\mathcal{A}$ is a finite directed graph with vertices called *states*, written as small circles; a special vertex called the *start state* of W, with the letter 's' (for 'start') written inside it; and directed edges connecting the vertices, each edge labelled with a letter from $\mathcal{A}$. For any given label, each vertex can have at most one edge directed out of it with this fixed label. Finally, we pick a subset Y of states which we call *accept states*, and draw the vertices of accept states as double circles. States not in Y are referred to as *fail states*. Examples of some finite state automata are given in figures 2 and 3.

A finite state automaton W gives a language over $\mathcal{A}$ as follows: If $w = w_o w_1 \cdots w_n$ is a word with each $w_i \in \mathcal{A}$, then beginning at the start state s, move from state to state by reading along the edge labelled w_0, then along $w_1, \ldots,$ then along w_n. If at any time the current state is v,

Figure 2: Some FSA and their accepted languages.

and the next letter which should be read is w_i but there is no edge directed out of v labelled w_i, then w is not accepted by W. If after reading the word w the current state is an accept state, then w is accepted by W, otherwise w is not accepted by W. The set of accepted words $L = L(W)$ is said to be the *language accepted by* W. Note that if the start state s is an accept state then the empty word is an element of L. A language which is accepted by a finite state automaton is called a *regular language.*

Regular languages are the simplest, most important languages studied by computer scientists. They are very special - most languages (i.e., subsets of $\mathcal{A}^*$) are not regular. An example of a language over the alphabet $\mathcal{A} = \{a, b\}$ which is not regular is the set $\{a^n b^n : n \in Z\}$; this follows immediately from the well-known "pumping lemma" of computer science (see [HU]). Note that $\{a^n b^m : n, m \in Z\}$ is regular. The reason why $\{a^n b^n : n \in Z\}$ is not regular is that a finite state automaton has no memory, and so cannot know exactly how many b's to accept after having accepted n a's.

There are other ways to define regular languages via certain grammatical operations and other machines similar to the automata described above (see [Ep1],[E et al.] or [HU]). Different (though equivalent) definitions of regular languages are useful in different situations, but for us the above definition will suffice.

Before giving the definition of automatic group we will need the notion of a *two-variable padded language.* Given an alphabet $\mathcal{A}$, we can add a padding symbol $\$ \notin \mathcal{A}$ to form the alphabet $\mathcal{A} \cup \{\$\}$, and we can consider a finite state automaton W as above, but this time with labels in $(\mathcal{A} \cup \$) \times (\mathcal{A} \cup \$) \backslash (\$, \$)$. Given a pair of words $(u, v) \in \mathcal{A}^* \times \mathcal{A}^*$, say $u = u_1 \cdots u_n, v = v_1 \cdots v_m$ with $m \leq n$, we pad v with the symbol $\$$ so that the resulting words have equal length. We will say that (u, v) is accepted by W if we can read off the edges $(u_1, v_1), \ldots, (u_m, v_m), (u_{m+1}, \$), \ldots, (u_n, \$)$ and end up at an accept state of W. The set of accepted pairs (u, v) is said to be *regular over the (padded) alphabet* $\mathcal{A}$. The point of padding is that pairs of words can be read at equal speeds, even if the words have different lengths.

4 Automatic Groups: Definitions and Examples

The definition of automatic group involves only finite state automata. We will later show this to be equivalent to a more geometric, and perhaps easier to understand, condition.

Let G be a group with finite generating set $\mathcal{A} = \{a_1, \ldots, a_n\}$ such that $\mathcal{A}$ actually generates G as a monoid. $\mathcal{A}$ is most often chosen as $\mathcal{A} = S \cup S^{-1}$, where S is a finite set of (group) generators for G and S^{-1} is the set of inverses of the elements of S. Notice that there is a natural map from $\mathcal{A}^*$, the free monoid on $\mathcal{A}$, to the group G which takes a word to the group element which it represents; we will denote this map by $w \mapsto \overline{w}$. G is an *automatic group* if the following conditions hold:

1. There is a language L over $\mathcal{A}$ given by a finite state automaton W such that the natural map $\pi : L \longrightarrow G$ is onto; the image $\pi(w)$ is denoted by $\overline{w}$. Thus there is at least one, and perhaps even an infinite number, of words in L representing each group element. W is called the *word acceptor*, and it gives a choice of canonical forms for group elements.

2. The following (padded) languages are regular:

$$\begin{aligned} L_= &= \{(u, v) : u, v \in L \text{ and } \overline{u} = \overline{v}\} \\ L_{a_1} &= \{(u, v) : u, v \in L \text{ and } \overline{u} = \overline{v a_1}\} \\ &\vdots \\ L_{a_n} &= \{(u, v) : u, v \in L \text{ and } \overline{u} = \overline{v a_n}\}. \end{aligned}$$

$L_=$ gives a FSA which checks whether two canonical forms represent the same group element; its accepting automaton $W_=$ is called the *equality checker.* L_{a_i} gives a FSA which checks whether two canonical forms represent group elements which differ by a_i (multiplied on the right); its accepting automaton W_{a_i} is called a *word comparator.*

The collection of automata $(W, W_=, W_{a_1}, \dots, W_{a_n})$ is called an *automatic structure* for G. One should think of the word acceptor W as a way of choosing canonical forms for group elements, and the other automata of the structure as a way of relating and piecing together these canonical forms to give the group. We begin with three immediate simple examples.

Simple examples of automatic groups :

1. Finite groups are automatic. If $G = \{g_1, \dots, g_n\}$ is a finite group of order n, we take the set $\mathcal{A} = \{g_1, \dots, g_n\}$ as monoid generating set, and let $L = \mathcal{A}$. Then L is a regular language since finite sets are regular (the automaton is an n-segment star with labels g_i on the edges). Since the sets $L_=, L_{g_i}, \dots, L_{g_n}$ are finite, hence regular, G is automatic.

2. The infinite cyclic group $Z =< a >$ is automatic. Figure 3 gives the automatic structure for Z with respect to the generating set $\mathcal{A} = \{a, A = a^{-1}\}$.

3. The free group on two generators $F(a, b)$ is automatic. The word acceptor given in figure 2, which accepts reduced words in $F(a, b)$, is part of an automatic structure. The reader is invited to construct the equality checker and the comparator automata.

An automatic group may have many automatic structures, even for a fixed generating set; the point is to find an automatic structure which is natural and easy to understand. It has been shown that if G is automatic with respect to one finite set of generators, then it is automatic with respect to any other finite set of generators ([E et al.]).

There is another definition of automatic group which is equivalent to the definition we have given. This second definition is more geometric than the first, and so it is often useful in proving that groups arising in geometric situations are automatic. We shall prove the equivalence of the two definitions since the proof gives a taste of the interplay between the geometry and the finite state automata.

Figure 3: Auotmatic structure for $Z =< a, A = a^{-1} >$.

Recall that a path u in $\Gamma_S(G)$ can be thought of as a map $u : [0, \infty) \longrightarrow \Gamma_S(G)$, where $u(t)$ is the element of G given by the first t letters of u if t is less than the length of u, and $u(t) = \overline{u}$ if t is greater than or equal to the length of u (recall that $\overline{u}$ is the element of G represented by the word u in the free group on S). Two paths u and v in $\Gamma_S(G)$ are said to satisfy the k-*fellow traveller property* if $d_{\Gamma_S(G)}(\overline{u(t)}, \overline{v(t)}) \leq k$ for all $t \geq 0$.

Proposition 1 *A group G is automatic if and only if the following properties hold:*

1. *G has a word acceptor W with regular language $L(W)$ over some finite monoid generating set $\mathcal{A}$, as in condition 1 of the definition of automatic group. Recall that the natural map $L(W) \longrightarrow G$ is required to be onto, and that the image of a word w is denoted by $\overline{w}$.*

2. *There is a constant k such that if $u, v \in L(W)$ represent elements of G which are distance 1 apart in $\Gamma_{\mathcal{A}}(G)$, then the paths u and v satisfy the k-fellow traveller property.*

Proof: If G is automatic with monoid generating set $\mathcal{A} = \{a_1, \ldots, a_n\}$, let c be an integer greater than the maximum number of states in any of the word comparators $W_{a_1}, \ldots, W_{a_n}$. Suppose $\overline{u}$ and $\overline{v}$ are a distance 1 in $\Gamma_{\mathcal{A}}(G)$, so that $u, v \in L$ differ by some a_i, say $u = va_i$. Let $s(t)$ denote the state W_{a_i} is in after reading the (possibly padded) prefixes $u(t)$ and $v(t)$. Then clearly there is a path in W_{a_i} from $s(t)$ to an accept state, and this path must have length less than c. Note that the "path to an accept state" in W_{a_i} gives a pair of paths in $\Gamma_{\mathcal{A}}(G)$ from $\overline{u(t)}$ and $\overline{v(t)}$ to points in $\Gamma_{\mathcal{A}}(G)$ which differ by the generator a_i; each of these paths must have lengths (in the Cayley graph) less than c. From this it follows easily that $d_{\Gamma_S(G)}(\overline{u(t)}, \overline{v(t)}) \leq 2(c-1)+1 = 2c-1$ for all t (see figure 4). This bound holds uniformly for all such paths.

Now suppose G satisfies conditions (1) and (2) of the hypothesis. We'll build a finite state automaton *Diff* which keeps track of how two paths differ over time, and we'll use *Diff* to build the W_{a_i}. If S and s_0 are the state set and start state of the word acceptor W, and if B_k is the set of group elements of distance $\leq k$ from the identity $e \in G$ (i.e. the k-ball in $\Gamma_S(G)$ around e), let the state set S' of *Diff* be $S \times S \times B_k$, and let the start state of *Diff* be (s_0, s_0, e). If *Diff* is in state (s_1, s_2, g) and the letter (x, y) in the associated padded alphabet is read, then *Diff* goes into the state $(t_1, t_2, x^{-1}gy)$, where x (resp. y) takes state s_1 (resp. s_2) of W to state t_1 (resp. t_2) of W (with the padding symbol not changing the states); except if t_1 or t_2 is a fail state of W or if $x^{-1}gy \notin B_k$, in which case *Diff* goes into fail state (don't draw an arrow labelled (x, y) coming out of (s_1, s_2, g)). Now *Diff* is a FSA which keeps track of whether two paths are accepted by W, and also keeps track of how far away the two paths are (see figure 5). The automata W_{a_i} may then be defined by taking the FSA *Diff* with accept states of the form (s_1, s_2, a_i), where $s_1, s_2 \in S$. $W_=$ is given by *Diff* with accept states of the from (s_1, s_2, e), where $s_1, s_2 \in S$.

$\diamond$

Proposition 1 shows that an automatic structure for a group G with generating set $\mathcal{A} = \mathcal{A}^{-1}$ is determined by the regular set $L \subseteq \mathcal{A}^*$; the word acceptor for L and the comparator automata must exist, but they need not be given explicitly. Most often one shows that a certain set of words L, such as the set of geodesics in the Cayley graph, is regular, and that the k-fellow traveller property is satisfied for some k. The proof of Proposition 1 also shows that the entire Cayley graph $\Gamma_{\mathcal{A}}(G)$ is determined by the k-ball around the identity; the "linear recursion" given by the automata of the automatic structure may be used to knit together copies of this k-ball to obtain $\Gamma_{\mathcal{A}}(G)$.

Figure 4: Proof of k-fellow traveller property for an automatic group.

Figure 5: The idea behind the FSA *Diff*.

Automatic groups encompass a large class of examples under one theory. We provide a list of the most well-known examples, others are being discovered at a rapid pace. The proof that a given class of groups is automatic usually involves doing quite a bit of geometry in spaces on which that class of groups acts in a geometric way (e.g., cocompactly by isometries); hence the proofs of the facts below have a strongly geometric flavor.

Main examples of automatic groups:

1. **Finite groups**.

2. **Abelian groups**.

3. **Negatively curved groups**. Negatively curved groups (sometimes called 'hyperbolic groups') are those groups whose cayley graphs have uniformly thin triangles. These groups have been studied extensively in the last ten years (see [Gr, Ca2, Gh, CDP, GdlH] for surveys). Examples of negatively curved groups include finite groups, free groups, cocompact discrete groups of hyperbolic isometries (more generally, fundamental groups of compact manifolds with strictly negative sectional curvatures), and small cancellation groups satisfying the usual metric small cancellation conditions ([LS]). Gromov has claimed that, in some combinatorial sense, "most" groups are negatively curved. The proof that negatively curved groups are automatic is essentially contained in Cannon's original paper ([Ca1]), although it is of course not couched in those terms; the language L of normal forms consists of the set of geodesic words. Negatively curved groups were the first and are still the most important examples of automatic groups. Automatic groups are much more general than negatively curved groups; for example, negatively curved groups cannot have a subgroup isomorphic to $Z \times Z$. In fact, any Seifert fiber space over a closed surface has an automatic fundamental group ([E et al.]), but most of these spaces do not even admit metrics of non-positive curvature.

4. **Non-metric small cancellation groups.** Groups with a presentation satisfying the weaker, non-metric small cancellation hypothesis are not, in general, negatively curved. However, Gersten and Short have shown that such groups are automatic ([GS1, GS2]). In some sense the theory of automatic groups unifies and supercedes small cancellation theory.

5. **Many Coxeter groups**. Many Coxeter groups are automatic. It still seems to be an open question whether all Coxeter groups are automatic. It is quite enjoyable to construct explicitly automatic structures for reflection groups in the euclidean and hyperbolic planes, such as the group of reflections in the sides of a right-angled pentagon in the hyperbolic plane; the reader is encouraged to do so.

6. **Most three-manifold groups**. The situation for compact three-manifolds is pretty well understood. Epstein and Thurston have shown ([E et al.]) that if M is a compact three-manifold which satisfies Thurston's Geometrization Conjecture, then $\pi_1(M)$ is automatic if and only if none of the pieces in its decomposition along spheres and tori is modelled on Nilgeometry or Solvgeometry. That is, if $M_1, \ldots, M_k$ are compact three-manifolds whose interiors are modelled on one of the eight three-dimensional geometries, and if M is the compact, connected three-manifold formed from the M_i by connected sum, disk sum and identifying boundary tori in pairs, then $\pi_1(M)$ is automatic if and only if none of the M_i is closed and modelled on Nilgeometry or Solvgeometry. Another approach to this, via theorems on the automaticity of graphs of groups, is given by Shapiro ([S]).

7. **Geometrically finite groups**. Epstein has shown ([E et al.]) that every geometrically finite hyperbolic group is automatic; in particular, fundamental groups of hyperbolic link complements are automatic. This is useful since most link complements have a hyperbolic structure. The main part of Epstein's proof involves figuring out what the quasi-geodesics (see below) are in the universal cover of a finite volume hyperbolic manifold with its cusps cut off.

8. **Braid groups**. Thurston has shown ([E et al.]) that the braid group on n strands is automatic (for each $n \geq 1$), which also shows that the mapping class group of the $(n+1)$-punctured sphere is automatic. This work explores several algorithmic aspects of the braid group. Thurston has conjectured that the mapping class groups of all hyperbolic surfaces are automatic.

The property of being automatic is closed under direct product, free product and free product with amalgamation over a finite subgroup. If H is a finite index subgroup in G, then H is automatic if and only if G is. These closure properties give many more examples of automatic groups;

in particular, cocompact discrete groups of Euclidean isometries are automatic since they contain abelian subgroups of finite index by Bieberbach's Theorem.

Although a wide variety of examples are automatic, this class of groups is very special, much more so than, say, groups with solvable word problem. To show that a group is not automatic seems difficult, for how does one show that "There does not exist any regular language such that ..."? However, techniques for showing that certain groups are not automatic have been developed; most of these involve isoperimetric inequalities in groups. We refer the reader to [E et al.],[Ge1], and [GS3] for details.

Examples of groups that are not automatic :

1. **Infinite torsion groups**. The mere existence of such groups is far from trivial, so it is not very disconcerting that such groups are not automatic (it is, perhaps, heartening). That infinite torsion groups are not automatic follows immediately from the well-known "pumping lemma" for finite state automata (see [Gi] for the proof).

2. **Nilpotent groups**. Finitely generated nilpotent groups which do not contain an abelian subgroup of finite index are not automatic. This was first proved by Holt. For example the three-dimensional Heisenberg group $H_3 =< a,b,c : [a,b] = c,[a,c] = 1 = [b,c] >$, the simplest non-abelian nilpotent group, has a cubic isoperimetric function (see property 7 below) and so is not automatic. The fact that nilpotent groups are not automatic is a bit surprising and annoying, considering the fact that nilpotent groups are quite common and have an easily solved word problem.

3. $\boldsymbol{SL_n(Z), n \geq 3}$. Note that $SL_2(Z)$ contains a free subgroup of index six, and so is automatic. The proof that $SL_n(Z), n \geq 3$ is not automatic involves finding a contractible manifold on which $SL_n(Z)$ acts with compact quotient, and showing that a higher-dimensional isoperimetric inequality is not satisfied by that space. The search for this manifold involves the study of the symmetric space $SL_n(R)/SO_n(R)$.

4. **Baumslag-Solitar Groups**. The group $G_{p,q} =< x,y : yx^py^{-1} = x^q >$ is not automatic unless $p = 0$, $q = 0$ or $p = \pm q$. These groups provide examples of groups which are not automatic but are *asynchronously automatic* (see [BGSS, E et al.]).

5 Hyperbolic Groups are Automatic

It is most often the case that proving that a group G is automatic requires doing quite a bit of geometry in a space on which G acts in a geometric way. As an example we prove the result of Cannon that cocompact discrete groups of hyperbolic isometries are automatic; in fact we show this more generally for fundamental groups of compact manifolds with (not necessarily constant) strictly negative sectional curvatures.

A path $\alpha : [a,b] \longrightarrow X$ in a metric space X is a *quasi-geodesic* if it is a geodesic up to constants; that is, there exists a K such that

$$1/K(t_2 - t_1) - K < d_X(\alpha(t_1), \alpha(t_2)) < K(t_2 - t_1) + K$$

for all $t_1 < t_2$ in $[a,b]$; in this case α is called a K-*quasi-geodesic*. A quasi-geodesic is a 'geodesic in-the-large' (hence the adding and subtracting of the constant K). One of the most important facts about negatively curved spaces is that quasi-geodesics are close to geodesics.

Lemma 2 (Morse-Mostow) *Let M be a compact manifold with strictly negative sectional curvatures. Then there exists a constant $C = C(K)$ such that any finite K-quasi-geodesic α in the universal cover $\widetilde{M}$ lies in the C-neighborhood of the geodesic joining the endpoints of α.*

This lemma, implicit in a 1924 paper of Morse [1] ([Mo]), was used in the proof of Mostow's rigidity theorem; a proof is given in [Th2]. Note that the situation is quite different for spaces which are not negatively curved; consider the logarithmic spiral in the euclidean plane.

If G is the fundamental group of a compact Riemannian manifold M, then a natural copy of the Cayley graph $\Gamma_S(G)$ sits inside of the universal cover $\widetilde{M}$ of M; namely, choose a basepoint $m \in M$ and a lift $\tilde{m} \in \widetilde{M}$ of m, put a vertex at $g(\tilde{m})$ for each deck transformation $g \in G$, and for each edge from g to $g \cdot s$ in the Cayley graph connect $g(\tilde{m})$ and $gs(\tilde{m})$ by a geodesic in $\widetilde{M}$ (here $\widetilde{M}$ is given the metric induced by that on M, so that the deck transformations act as isometries). It is a fundamental fact that paths in the Cayley graph $\Gamma_S(G)$ that are geodesic are actually quasi-geodesics in $\widetilde{M}$. This ties the (quasi-)geometry of the fundamental group together with the (quasi-)geometry of the universal cover.

[1] A. Haefliger pointed this out to me.

Figure 6: Why geodesics in $\Gamma_S(G)$ are fellow travellers.

Theorem 3 *If M is a compact manifold with strictly negative sectional curvatures, and if S is any generating set for $G = \pi_1(M)$, then the set of geodesic (shortest) words in $\Gamma_S(G)$ is a regular language, and is part of an automatic structure for G; in particular G is automatic.*

Proof: We shall prove that the set L of geodesics in $\Gamma_S(G)$ is a regular language satisfying the k-fellow traveller property for some k.

By the comment above there is some constant K such that geodesics $u, v \in L$ which represent elements of G at distance one apart in $\Gamma_S(G)$ are K-quasi-geodesics in $\widetilde{M}$, so by Mostow's Lemma they lie in a C-neighborhood of geodesics u' and v' in $\widetilde{M}$ with the same endpoints as u and v; this is the only place where the strictly negative curvature assumption is used. Now u' and v' form two sides of a triangle whose third side has length at most $2K$, by the equation on page 14. Since $\widetilde{M}$ is a non-positively curved space, u' and v' are within Hausdorff distance $2K$ from each other. Hence u and v are Hausdorff distance of at most $2K + C + C$ apart (see figure 6),

and from this it follows easily that, since u and v are geodesics in $\Gamma_S(G)$, they are k-fellow travellers for some constant k, and this k depends only on the curvature bound on M, not on u and v.

Notice that the subwords u_i and v_i consisting of the first i letters of u and v are geodesics, and $u_i v_i^{-1}$ lies in the k-ball B_k around the identity in $\Gamma_S(G)$.

Figure 7: How to determine the new state after being in state T and reading the generator g. Here $t_1, t_2 \in T, a \in S \cup S^{-1}$.

We now build a finite state automaton W which recognizes precisely the set of geodesics in $\Gamma_S(G)$. As state set S of W we take the set of subsets of the k-ball B_k together with a fail state F, let the subset consisting of the identity be the start state, and take every state of S except for F as an accept state. Suppose the generator g of the word w is read when W is in the state T. Then W should go into the fail state if either T is the fail state or $g \in T$; otherwise W should go into the state $\{g^{-1}ta : t \in T, a \in S \cup S^{-1} \cup \{e\}\} \cap B_k$ (see figure 7). The idea is that, after reading in the first i letters of w, the current state S_i knows all other possible shortest words of length i which are subwords of words representing $\overline{w}$ (remember that these words must differ from w_i by at most an element of B_k). We can keep track of all of these words with finitely states since we are really only keeping track of word differences, which lie in a finite set (namely B_k). If the $(i+1)$'st generator of w lies in S_i, then w_{i+1} (hence w) is not a geodesic, since $\overline{w_{i+1}}$ can be represented by a word of length i. Conversely, if w is not a geodesic, then we will enter the fail state upon reading the smallest initial subword of w which is not a geodesic. This shows that W accepts precisely the set of shortest words; hence L is regular. $\diamond$

Cannon's original proof of theorem 3 is based on the notion of 'cone type' (see [Ca1, Ep1]). The idea is that geodesics in a hyperbolic group have only finitely many asymptotic behaviors (i.e. cone types), so building a geodesic generator by generator requires looking at a finite set (the set of cone types) and applying finitely many rules (determining the possible cone types after adding the next generator).

6 Interesting Properties

The class of automatic groups encompasses a wide variety of groups, many arising from vastly different geometric situations and exhibiting vastly different behaviors. It is indeed surprising that such a diverse class of groups should share all of the properties which come from being automatic.

Some nice properties of automatic groups :

1. **Finite presentation**. Although it is true *a priori* only that automatic groups are finitely generated, it is in fact true that every automatic group has a finite presentation; that is, a presentation with finitely many generators and finitely many relators.

2. **Fast solution to word problem**. There is a quadratic time algorithm to solve the word problem for an automatic group; quadratic in the sense that one can check in $O(n^2)$ steps whether a word of length n represents the identity or not. In fact, one can put a word in a canonical form (given by the regular language of the word acceptor) in quadratic time([E et al.]). We should also mention that, if G is automatic, then there is some automatic structure for which there is a unique canonical form for each group element; that is, the natural map $\pi : L \longrightarrow G$ is bijective. It is not known if automatic groups have solvable conjugacy problem.

3. **Fast pictures**. One can enumerate the elements of an automatic group with a unique word for each element; the first n elements can be enumerated in time $O(n \log n)$. This allows for efficient construction of the Cayley graph of an automatic group, as well as pictures of sets which are invariant under the action of an automatic group. Mumford and Wright have used automatic groups in programs to draw pictures of limit sets of quasifuchsian groups efficiently; mathematicians have been trying to draw such pictures efficiently on computers for many years. Automatic group structures have also been used in

relating frames in the soon to be released movie "Not Knot". After implementing ideas from automatic groups, the time for computing the numbers used to locate points in a single frame went from about 20 minutes to about 15 seconds on a four processor Iris. Since there are 28 frames per second, the amount of time saved is quite substantial ([Ep2]).

4. **Uniform algorithms**. There is an algorithm which takes as input a finite presentation and as output gives the automatic structure $(W, W_=, W_{a_1}, \ldots, W_{a_n})$ for the group. The algorithm does not terminate if the group is not automatic; in fact, there is no algorithm which can determine for a presentation $G =< x_1, \ldots, x_n : r_1, \ldots r_m >$ of a group G whether or not the G is automatic. The algorithm exists since it is possible to give a list of (checkable) axioms characterizing when a collection of finite state automata form the automatic structure of a group. The algorithm for solving the word problem for a fixed automatic group is itself algorithmically constructible, so there is a uniform algorithm for solving the word problem for all automatic groups! It should be noted that the algorithm which finds the automatic structure from the presentation is extremely slow. More efficient methods for finding automatic structures in slightly more special cases have in fact been programmed by Epstein, Holt and Rees. Their ideas involve the Knuth-Bendix process, and their programs have been quite successful at finding automatic structures for a number of examples ([EHR]). One can also show that there is an algorithm which takes as input a presentation of an (*a priori*) automatic group, and as output tells whether or not the group is trivial, and whether or not the group is finite. For presentations in the class of arbitrary groups, these problems have been shown to be unsolvable ([Ra]). It is an open question whether the isomorphism problem is solvable for automatic groups; that is, if there is an algorithm to determine, given two presentations each of whose groups is automatic, whether or not the groups are isomorphic.

5. **Rational growth functions**. Many automatic groups have rational growth functions; in particular those groups which have an automatic structure where the language of accepted words consists of geodesics have rational growth functions. Recall that if we are given a presentation (G, S), and if c_n denotes the number of elements of $\Gamma_S(G)$

at distance n from the identity, the counting function for (G, S) is the function $f(x) = \sum_{i=1}^{\infty} c_n x^n$. The rationality of $f(x)$ may be interpreted as the fact that the number of elements contained in the sphere (or ball) of radius n in $\Gamma_S(G)$ may be determined by a linear recursion, such as the recursion which gives the Fibonacci sequence. The automatic groups team at Warwick has programs which, given the automatic structure of an automatic group where the language of accepted words consists of geodesics, produces a rational function giving the growth of the group. It is an open question whether all automatic groups have rational growth functions.

6. **Type FP_∞.** If G is an automatic group, then there exists an Eilenberg-MacLane space $K(G, 1)$ with finitely many cells in each dimension; in this case G is said to be of type FP_∞ (see [Al]). It still seems to be unknown whether torsion-free automatic groups must have finite cohomological dimension.

7. **Quadratic isoperimetric function.** We shall not discuss isoperimetric functions in groups here; the reader may consult [Ge1, Sh, E et al.]. Isoperimetric functions in groups are extremeley interesting, and have become quite important in combinatorial group theory and geometry; related concepts have recently proven useful in the study of three-manifolds ([Ge2, St]). Gromov showed that the negatively curved groups are precisely those which have a linear isoperimetric function. Automatic groups satisfy a quadratic isoperimetric function, but are not characterized by this property. Thurston (unpublished) has shown that the five-dimensional Heisenberg group has a quadratic isoperimetric function but is not automatic.

It is possible to get a feel for the breadth and unifying power of the theory of automatic groups by matching up groups in the list of examples with properties from the list just given. For instance, automatic groups give a uniform quadratic time solution to the word problem for fundamental groups of compact negatively curved manifolds, most Coxeter groups, and the braid groups (previously known algorithms for the braid groups never discussed speed, and seem to be much slower). The reader may wish to contemplate other "theorems" obtained by matching pairs in the lists.

7 Related Topics, Open Problems, and a Vision of the Future

The field of automatic groups (and related topics) is still quite young; accordingly, there are many open questions which are interesting, easy to state, and perhaps not so difficult for a newcomer to think about. Listed below are a few personal favorites. For other open questions, the reader is encouraged to dive into the references given at the end of this paper, in particular [Ge3].

Some open problems :

1. Prove that the mapping class groups of hyperbolic surfaces are automatic. As a (perhaps) easier question, show that these groups satisfy a quadratic isoperimetric inequality ([Ge1, E et al.]).

2. Are cocompact lattices in $SL_3(R)$ automatic? Note that $SL_3(Z)$ is a lattice in $SL_3(R)$ which is not cocompact and not automatic. There is a p-adic analog to this question which has been solved ([GS1]). Find examples of other arithmetic groups which are or are not automatic. So far not much seems to be known for arithmetic groups, except for a result of Gersten and Short ([GS3]) which shows that $SL_2(\mathcal{O})$, with $\mathcal{O}$ a real quadratic number field, is not biautomatic (see below). Note that for some rings of algebraic integers, such as $\mathcal{O} = Z[i]$, the group $SL_2(\mathcal{O})$ is the fundamental group of a three manifold which is automatic.

3. Are fundamental groups of compact, non-positively curved manifolds automatic? The answer to this question is probably "no" (see 9 below).

4. A simple question whose answer has eluded everyone: If $G \times H$ is automatic, is G automatic? The corresponding statement for free products is true ([BGSS]).

5. Negatively curved groups have a well-defined "boundary at infinity" which gives a great deal of information about such groups (see any of the surveys cited above). Is there a corresponding theory for automatic groups?

6. Explore the effect of changing generators and automatic structures on the constant k of the k-fellow traveller property. What is the best constant you can get for a specific example (e.g. a surface group)?

For a given group G, what is the minimal number of states of a word acceptor which is part of an automatic structure with unique representatives for G?

7. Study the quasi-convex subgroups of automatic groups (see [GS3]), as well as other geometric properties which have algorithmic consequences. Work this out explicitly for fundamental groups of three-manifolds.

8. Does every automatic group have a rational counting function? This is true for automatic groups where the language of accepted words consists of geodesics. Explore analytic properties of these functions in special cases (Cannon and others have done this for several examples).

9. A group is *combable* if there is a section $\sigma : G \to \mathcal{A}^*$ of the natural map $\pi : \mathcal{A}^* \to G$ which satisfies the k-fellow traveller property for some k for all paths. Automatic groups are simply combable groups whose image $\sigma(G) \subseteq \mathcal{A}^*$ is a regular language. Much of the theory of automatic groups has been generalized to combable groups ([Sh]). Combability is also a very natural condition to look at when studying geometry, particularly the geometry of nonpositive curvature. Find an example of a combable group which is not automatic. A good place to look (according to Thurston) might be at a cocompact group of isometries of $H^2 \times H^2$ which does not have a product of surface groups as a subgroup of finite index. This would also show that the property of being automatic is not a so-called "geometric invariant", i.e., a quasi-isometry invariant, but depends on more combinatorial properties of the group.

10. Are automatic groups residually finite? Gersten has recently found a combable group that is not residually finite. Which automatic groups admit a faithful linear representation (such groups are residually finite)? Try to construct interesting representations of automatic groups using the automatic structure.

11. A notion stronger than automatic is that of *biautomatic*, where one is also supplied with word comparators W'_{a_i} for multiplication by a_i on the left. Much more is proven about biautomatic groups than automatic groups (e.g. biautomatic groups have solvable conjugacy problem), although it is not known whether every automatic group

is biautomatic. There seems to be a deep theory of subgroup structure for biautomatic groups, as has been developed by Gersten and Short ([GS3]). Carry over the theory of biautomatic groups to automatic groups, in particular solve the conjugacy problem for automatic groups and also find theorems which put some constraint on what the subgroups of an automatic group can be (see [Ge1, GS3]). Even better, determine whether every automatic group is in fact biautomatic.

12. Generalize the entire theory of automatic groups by using machines which are more complicated than finite state automata (see below).

Thurston has envisioned a program of studying algorithmically groups that arise naturally in geometry and topology. Automatic groups are the first stage of this program. One idea is to relativize the theory by replacing the states of a finite state automaton by black boxes which could do computations in nilpotent groups in order to study groups which are automatic (or hyperbolic) "relative to" certain nilpotent subgroups; examples being non-cocompact lattices in Lie groups of R-rank 1 and fundamental groups of nonpositively curved manifolds. Another direction might be to replace finite state automata by more complicated machines. In the hierarchies of languages and machines studied by complexity theorists (e.g. the Chomsky Hierarchy), regular languages and finite state automata are always at the bottom of the ladder; in fact, regular languages may be characterized by the fact that it takes zero-space on an (off-line) Turing machine to recognize them (see [HU]). More complicated machines should allow us to do computations (such as solving the word problem efficiently) in more complicated groups; the geometry of the group dictating the nature of the machine. The possibilities seem limitless.

References

[Al] J. Alonso, 'Combings of Groups', in *Proceedings of the Workshop on Algorithmic Problems*, C.F. Miller III and G. Baumslag, ed., Springer MSRI series, 1990 (to appear).

[BGSS] G. Baumslag, S.M. Gersten, M. Shapiro and H. Short, 'Automatic Groups and Amalgams', OSU Mathematical Research Institute Preprints, 1990 (to appear in *Journal of Pure and Applied Algebra*).

[Ca1] J. Cannon, 'The Combinatorial Structure of Cocompact Discrete Hyperbolic Groups', *Geometriae Dedicata* **16** (1984), 123-148.

[Ca2] J. Cannon, 'The theory of negatively curved spaces and groups', in *Ergodic theory, symbolic dynamics, and hyperbolic spaces*, edited by T. Bedford, M. Keane, C. Series, Oxford Univ. Press, 1991.

[CDP] M. Coornaert, T. Delzant, A. Papadopoulos, *Notes sur les groupes hyperboliques de Gromov*, Springer Lecture Notes, vol. 1441.

[E et al.] D.B.A. Epstein, J. Cannon, D.F. Holt, S. Levy, M.S. Paterson and W.P. Thurston, *Word Processing and Group Theory*, preprint, Fall 1991 (to be published).

[EHR] D.B.A. Epstein, D.F. Holt, and S.E. Rees, 'The use of Knuth-Bendix methods to solve the word problem in automatic groups', *Journal of Symbolic Computation*, 1991.

[Ep1] D.B.A. Epstein, 'Computers, Groups and Hyperbolic Geometry', *Astérisque* 163-164 (1988), pp. 9-29.

[Ep2] D.B.A. Epstein, personal communication.

[Ge1] S.M. Gersten, 'Dehn Functions and ℓ_1-norms of finite presentations', in *Proceedings of the Workshop on Algorithmic Problems*, C.F. Miller III and G. Baumslag, ed., Springer MSRI series, 1990 (to appear).

[Ge2] S.M. Gersten, 'The linear isodiametric inequality for groups and three-manifolds', preprint.

[Ge3] S.M. Gersten, 'Problems on Automatic Groups', in *Proceedings of the Workshop on Algorithmic Problems*, C.F. Miller III and G. Baumslag, ed., Springer MSRI series, 1990 (to appear).

[GdlH] E. Ghys and P. de la Harpe, editors, *Sur les groupes Hyperboliques d'aprés Mikhael Gromov*, Progress in Mathematics, Vol. 83, Birkhaüser, 1990.

[Gh] E. Ghys, 'Les Groups Hyperboliques', Séminaire Bourbaki, march 1990.

[Gi] R. Gilman, 'Groups with a rational cross-section', in *Combinatorial Group Theory and Topology*, Annals of Math. study 111, ed. by S. Gersten and J. Stallings, Princeton Univ. Press, 1987.

[Gr] M. Gromov, 'Hyperbolic groups', *Essays in Group Theory*, Springer Verlag, MSRI series Vol. 8, S.M. Gersten, ed., 1987, pp. 75-263.

[GS1] S.M. Gersten and H. Short, 'Small cancellation theory and automatic groups', *Invent. math.* 102, 305-334 (1990).

[GS2] S.M. Gersten and H. Short, 'Small Cancellation Theory and Automatic Groups: Part II', MSRI preprint, July 1989, to appear in *Invent. Math.*

[GS3] S.M. Gersten and H. Short, 'Rational Subgroups of Biautomatic Groups', *Annals of Math.*, 134 (1991), 125-158.

[HU] J.E. Hopcroft and J.D. Ullman, *Introduction to Automata Theory, Languages and Computation*, Addison-Wesley 1979.

[LS] R.C. Lyndon and P.E. Schupp, *Combinatorial Group Theory*, Berlin-Heidelberg-New York: Springer 1977.

[Mo] M. Morse, 'A Fundamental Class of Geodesics on any Closed Surface of Genus Greater Than One', *T.A.M.S.*, Vol. 26, No. 1, 1924.

[Ra] M. Rabin, 'Recursive Unsolvability of Group Theoretic Problems', *Annals of Math.*, Vol. 67, No.1, Jan. 1958.

[S] M. Shapiro, 'Automatic Structure and Graphs of Groups", Ohio State Mathematical Research Institute Preprints, 1991.

[Sh] H. Short, 'Groups and combings', preprint, *Ecole Normale Supérieure de* LYON, July 1990.

[St] J. Stallings, 'Casson's idea about three-manifolds whose universal cover is R^3', preprint, July 1990.

[Th1] W.P. Thurston, 'Groups, tilings, and finite state automata', summer 1989 AMS Colloquium lectures, preliminary version.

[Th2] W.P. Thurston, *The Geometry and Topology of Three-Manifolds*, Princeton University Notes.

Princeton University

Current Address: Univ. of California, Berkeley
E-mail: farb@math.berkeley.edu