

# Four heads are better than three\*

Ville Salo<sup>1</sup>[0000–0002–2059–194X]

University of Turku

**Abstract.** We construct recursively-presented finitely-generated torsion groups which have bounded torsion and whose word problem is conjunctive equivalent (in particular positive and Turing equivalent) to a given recursively enumerable set. These groups can be interpreted as groups of finite state machines or as subgroups of topological full groups, on effective subshifts over other torsion groups. We define a recursion-theoretic property of a set of natural numbers, called unpredictability. It roughly states that a Turing machine can enumerate numbers such that every Turing machine occasionally incorrectly guesses (by either halting or not) whether they are in the set, even given an oracle for a prefix of the set. We prove that unpredictable recursively enumerable sets exist. Combining these constructions and slightly adapting a result of [Salo and Törmä, 2017], we obtain that four-headed group-walking finite-state automata can define strictly more subshifts than three-headed automata on a group containing a copy of the integers, confirming a conjecture of [Salo and Törmä, 2017]. These are the first examples of groups where four heads are better than three, and they show the maximal height of a finite head hierarchy is indeed four.

**Keywords:** torsion groups · finite-state automata · subshifts · recursion theory.

## 1 Introduction

In this paper, we combine construction techniques from group theory and recursion theory to verify a conjecture the author and Törmä made in [15] about group-walking automata. The a priori motivation was this conjecture, but the constructions and definitions we give may be of independent interest.

### 1.1 The results

**Theorem 1.** *Let  $A$  be  $\Sigma_1^0$ . Then there exists a recursively presented torsion group with bounded torsion, whose word problem is Turing equivalent to  $A$ .*

The exact types of reduction are that  $A$  many-one reduces to the word problem, and the word problem conjunctively reduces to  $A$ .

---

\* Research supported by Academy of Finland grant 2608073211.

Our recursion-theoretic contributions are of a somewhat technical nature (though not particularly difficult). We define a possibly new notion called  $\phi$ -impredictability (Definition 1), which roughly states that a Turing machine can enumerate numbers such that every Turing machine occasionally incorrectly guesses (by either halting or not) whether they are in the set, even given an oracle for a prefix of the set.

**Theorem 2.** *For every total recursive function  $\phi$ , there exists a  $\phi$ -impredictable  $\Sigma_1^0$  set.*

We now state our main ‘‘navigation-theoretic’’ contribution. In [15], for a f.g. (finitely-generated) group  $G$  and  $n \in N$ , the class of  $G$ -subshifts defined by group-walking finite-state machines with  $n$  heads is denoted by  $\mathcal{S}(G, n)$ . See [15] or Section 5 for the definitions. The following is a slight adaptation of a result of [15].

**Lemma 1.** *If  $G$  has bounded torsion and has  $\phi$ -impredictable word problem for fast-enough growing  $\phi$ , then  $\mathcal{S}(G \times \mathbb{Z}, 3) \subsetneq \mathcal{S}(G \times \mathbb{Z}, 4)$ .*

For example  $\phi = \exp \circ \exp \circ \exp \circ \exp \circ \exp$  is fast enough. By putting the above results together, we obtain that four heads are better than three, as claimed in the title.

**Theorem 3.** *There exists a finitely-generated recursively-presented group  $G$  containing a copy of the integers, such that  $\mathcal{S}(G, 3) \subsetneq \mathcal{S}(G, 4)$ .*

See Section 1.3 for some context for this result, and a new conjecture.

## 1.2 Some relevant existing work

The main group-construction result (recursively presented groups with bounded torsion and with a word problem of a prescribed difficulty) uses the idea from [13, 3] of groups of finite-state machines. It also uses existing torsion groups as a black-box, in particular it is an application of the deep theory that arose from the Burnside problem [1].

The idea of hiding information into the word problem is of course not a new idea in combinatorial group theory, but we are not aware of it appearing previously in the context of torsion groups. The Dehn monsters from [11], recursively presented groups where no infinite set of distinct elements can be enumerated, seem strongly related. (Our construction cannot be used to produce such groups due to using another group as a black box, but it is possible that their construction can be adapted to produce our result.)

Topological full groups (already on  $\mathbb{Z}$ ) are a well-known source of interesting examples of groups [10, 6, 9], and our groups can also be interpreted as subgroups of topological full groups of subshifts on torsion groups. In a symbolic dynamics context, [7] (independently) uses a similar construction to prove that the automorphism groups of multidimensional SFTs can have undecidable word problem.

Dan Turetsky showed in the MathOverflow answer [12] that the halting problem is not  $\phi$ -impredictable for some total recursive  $\phi$ , and that there exists a  $\Sigma_1^0$  set which is simultaneously  $\phi$ -impredictable for all total recursive  $\phi$ . The latter result of course implies Theorem 2.

### 1.3 Head hierarchies

For context, we state what is now known about the head hierarchies, and state a bold conjecture. See Definition 3 for the definitions.

**Theorem 4.** *For an infinite finitely-generated group  $G$ , denote by  $h(G) \in \mathbb{N}_{>0} \cup \{\infty\}$  the supremum of  $n$  such that  $\mathcal{S}(G, n-1) \subsetneq \mathcal{S}(G, n)$ . Then*

$$\{3, 4, \infty\} \subset \mathcal{H} = \{h(G) \mid G \text{ an infinite f.g. group}\} \subset \{2, 3, 4, \infty\}.$$

Our contribution is the 4 on the left, which is notable because the maximal finite height is now known.

*Proof (Proof sketch).* The fact  $4 \in \mathcal{H}$  follows from Theorem 3. The fact  $h \in \mathcal{H} \cap \mathbb{N} \implies h \leq 4$  is proved in [15], and the facts  $3 \in \mathcal{H}$  and  $\infty \in \mathcal{H}$  are proved in [15] and [14] respectively. The fact  $1 \notin \mathcal{H}$  is seen as follows: In [14] it is shown that  $\mathcal{S}(\mathbb{Z}, 1) \subsetneq \mathcal{S}(\mathbb{Z}, 2)$  and the proof easily generalizes to virtually  $\mathbb{Z}$  groups. On the other hand it is easy to see that the sunny-side-up  $X_{\leq 1}$  separates one from four heads, i.e.  $X_{\leq 1} \in \mathcal{S}(G, 4) \setminus \mathcal{S}(G, 1)$ , for all f.g. infinite groups  $G$  which are not virtually  $\mathbb{Z}$ , which can be proved using methods of [15] and [14].

We conjecture that four heads are needed if and only if the word problem is undecidable:

*Conjecture 1.* Let  $G$  be a finitely-generated infinite group which is not torsion. Then

- if  $\mathbb{Z} \leq G$  and  $G$  has decidable word problem, then  $h(G) = 3$ ,
- if  $\mathbb{Z} \leq G$  and  $G$  has undecidable word problem, then  $h(G) = 4$ .

The upper bounds are known, and  $h(G) = \infty$  is known for torsion groups [15]. The first item is the join of conjectures in [15] and [14]. It is known that the sunny-side-up subshift  $X_{\leq 1}$  does not prove  $h(\mathbb{Z}) = h(\mathbb{Z}^2) = 3$ , as on these groups two heads suffice for it. We do not know any other non-torsion groups where the sunny-side-up requires fewer heads than the above conjecture would suggest.

Settling this conjecture in the positive would not be the end of the story. A more refined invariant than  $h(G)$  would be to ask what the precise set of  $n$  such that  $\mathcal{S}(G \times \mathbb{Z}, n-1) \subsetneq \mathcal{S}(G \times \mathbb{Z}, n)$  is. In particular this is of interest when  $G$  is a torsion group; [15] only gives an affine function  $f(n) \sim 3n$  such that  $\mathcal{S}(G \times \mathbb{Z}, n) \subsetneq \mathcal{S}(G \times \mathbb{Z}, f(n))$ .

These results are of course about just one way to associate subshift classes to  $n$ -headed automata. We believe the results are relatively robust to changes in the definition, but some details are critical. It is in particular open what happens when  $G$  is a torsion group and the heads are allowed to communicate over distances, i.e. if they have a shared state.

## 2 Preliminaries

For two functions  $f, g$  write  $f = O(g)$  (resp.  $f = \Omega(g)$ ) if for some choice of  $a > 0$ ,  $f(n) \leq ag(n)$  (resp.  $f(n) \geq ag(n)$ ) for large enough  $n$ . Write  $f = \Theta(g)$  if  $f = O(g)$  and  $f = \Omega(g)$ . Write  $f \sim g$  for  $f(n)/g(n) \rightarrow 1$ . For  $a, b \in \mathbb{N}$  write  $a \mid b$  for “ $a$  divides  $b$ ”.

We assume some familiarity with computability/recursion theory, but we state some (not necessarily standard) conventions. We identify partial computable functions with Turing machines, and also their Gödel numbers. Let  $\mathcal{T} \subset \mathbb{N}$  be the set of codings of Turing machines which halt on every input, and write  $\mathcal{P} = \mathbb{N}$  for all partial recursive functions. “Recursive” means the same as “computable” and refers to the existence of a Turing machine, which always halts unless the function is explicitly stated to be partial. If  $\chi$  is a Turing machine, we write  $\chi(p) \downarrow$  if  $\chi$  halts on input  $p$ . A partial (not necessarily computable) function from  $A$  to  $B$  is denoted  $g : A \nrightarrow B$ .

Let us recall some basic definitions of reductions. A set  $A$  *many-one reduces* to a set  $B$  if there is  $\phi \in \mathcal{T}$  such that  $n \in A \iff \phi(n) \in B$ . For the following three reductions, we give quantitative versions, so we have a handle on the rate at which reduction happens. The definitions are stated in terms of characteristic sequences, but correspond to their usual meaning.

We say  $B$  *weakly truth table reduces*, or *wtt-reduces*, to  $A$  if there exists  $g \in \mathcal{P}$  and nondecreasing  $\beta \in \mathcal{T}$  with  $\beta(n) \rightarrow \infty$  such that when applied to words as  $g : \{0, 1\}^* \nrightarrow \{0, 1\}^*$ , if  $x \in \{0, 1\}^{\mathbb{N}}$  is the characteristic function of  $A$  and  $y$  that of  $B$ , we have that  $g(w)$  is defined on prefixes of  $y$ , and

$$w \prec y \implies (g(w) \prec x \wedge |g(w)| \geq \beta(|w|)).$$

We call  $\beta$  the *rate* of the reduction.

We say  $B$  *positively reduces* to  $A$  if it wtt-reduces to  $A$ , for some  $\beta \in \mathcal{T}$ , with the following additional properties for  $g$ :  $g \in \mathcal{T}$ ,  $|g(w)| = \beta(|w|)$  for all  $w$ , and  $g$  is *monotone* in the sense that  $u \leq v \implies g(u) \leq g(v)$ , where  $\leq$  is letterwise comparison. If further  $g(u)_i$  only depends on whether  $u \geq w_{i,|u|}$  for some  $w_{i,|u|}$  computable from  $i, |u|$ , then  $B$  *conjunctively reduces* to  $A$ .

We say  $B$  *Turing reduces* to  $A$  if there is an  $A$ -oracle machine that can determine membership in  $B$ . Clearly many-one reducibility implies conjunctive reducibility implies positive reducibility implies weak truth table reducibility implies Turing reducibility.

We assume some familiarity with group theory, but state some conventions. Our groups are discrete and mostly finitely-generated. Finitely-generated groups come with a finite generating set, which we usually do not mention. The identity of a group  $G$  is denoted by  $e_G$  (or just  $e$ ). For words  $u, v \in S^*$  where  $S$  is the generating set, write  $u \approx_G v$  (or just  $u \approx v$  if  $G$  is clear from context) when  $u$  and  $v$  represent (evaluate to) the same element of  $G$ . For  $g, h \in G$ , write  $[g, h] = g^{-1}h^{-1}gh$  for the *commutator* of  $g$  and  $h$ .

The *word problem* of a group  $G$  is the following subset  $W \subset \mathbb{N}$ : Let  $S$  be the fixed symmetric generating set, and order elements of  $S^*$  (finite words over

$S$ ) first by length and then lexicographically. Include  $n \in W$  if the  $n$ th word evaluates to the identity of  $G$ . Slight inconvenience is caused by linearizing the word problem this way, but on the other hand sticking to subsets of  $\mathbb{N}$  slightly simplifies the discussion in Section 3. We denote the word problem of  $G$  as  $\text{WP}(G) \subset \mathbb{N}$ .

If a countable group  $G$  acts on a compact zero-dimensional space  $X$ , the corresponding *topological full group* is the smallest group of homeomorphisms  $g : X \rightarrow X$  which contains every homeomorphism  $g$  with the following property: there exists a clopen partition  $(P_i)_{i=1}^k$  of  $X$  and  $g_i \in G$  such that for  $\forall i \in \{1, \dots, k\} : \forall x \in P_i : gx = g_i x$ . It turns out that the group contains precisely such homeomorphisms, i.e. they are closed under composition. One may think of  $((P_i)_i, (g_i)_i)$  as a local rule for  $g$ .

A group is *torsion* if all elements have finite order, i.e.  $G$  does not contain a copy of the integers. The *torsion function* of a torsion group is  $T_G : \mathbb{N} \rightarrow \mathbb{N}$  defined by

$$T_G(n) = \sup\{\text{ord}(g) \mid g \in G, |g| \leq n\}$$

where  $\text{ord}(g) = |\langle g \rangle|$  and  $|g|$  is the word norm with respect to the implicit generating set. A group is of *bounded torsion* if  $T_G(n) = O(1)$ .

In [8] Ivanov shows that the *free Burnside group*

$$B(2, 2^{48}) = \langle a, b \mid \forall w \in \{a, b, a^{-1}, b^{-1}\}^* : w^{2^{48}} \approx e \rangle$$

is infinite and has decidable word problem ([8, Theorem A]), and we obtain

**Lemma 2.** *There exists a finitely-generated torsion group with bounded torsion and decidable word problem.*

The decision algorithm is given in [8, Lemma 21.1]. For a survey on the Burnside problem see [1]. Theorem 3 could be proved using any torsion group with recursive torsion function (e.g. the Grigorchuk group), with minor modifications, as explained after the proof of Theorem 3. Our technical construction results (in particular Theorem 6) are stated and proved without assuming the existence of infinite f.g. groups of bounded torsion, though obviously to obtain Theorem 1 the existence of one is necessary (since it states the existence of one).

We assume some familiarity with symbolic dynamics on groups (see [2] for more information), but state some conventions. If  $G$  is a group, and  $\Sigma$  a finite set,  $\Sigma^G$  with the product topology and the  $G$ -action  $gx_h = x_{g^{-1}h}$  is called the *full shift*, and the actions  $(g, x) \mapsto gx$  are called *shifts*. A *subshift* is a topologically closed subset  $X$  satisfying  $GX = X$ . A particularly important subshift is the *sunny-side-up* subshift  $X_{\leq 1} = \{x \in \{0, 1\}^G \mid \sum_{g \in G} x_g \leq 1\}$ . Equivalently a subshift is defined by a family of *forbidden patterns*, i.e. a (possibly infinite) family of clopen sets that the *orbit*  $Gx$  of  $x$  may not intersect. If  $G$  has decidable word problem then  $X \subset \Sigma^G$  is *effective* if there exists a Turing machine that enumerates a family of forbidden patterns which defines the subshift. A *cellular automaton* on a subshift is a continuous shift-commuting self-map of it.

### 3 Impredictability

**Definition 1.** For a function  $\phi$ , a set  $A \subset \mathbb{N}$  is  $\phi$ -impredictable if

$$\exists \psi \in \mathcal{T} : \forall \chi \in \mathcal{P} : \exists^\infty p : \psi(p) \in A \iff \chi(p, A \upharpoonright \phi(p)) \downarrow.$$

To unravel this definition a bit, we want to have a Turing machine  $\psi$  which always halts and gives us positions  $\psi(p)$  on the number line  $\mathbb{N}$  so that, for any Turing machine  $\chi$ ,  $\psi(p) \in \mathbb{N}$  is just the halting information about  $\chi$  on input  $p$ , even if  $\chi$  is allowed oracle access to the first  $\phi(p)$  bits of  $A$ . Since  $\chi$  is quantified universally, this definition is one way to formalize the idea that it is hard to predict whether  $\psi(p) \in A$  even given access to the first  $\phi(p)$  bits of  $A$ .

**Lemma 3.** Let  $\phi', \phi \in \mathcal{T}$  and  $\phi'(n) \geq \phi(n)$  for all large enough  $n$ . Then every  $\phi'$ -impredictable set is  $\phi$ -impredictable, with the same choice of  $\psi$ .

*Proof.* Let  $\phi, \phi' \in \mathcal{T}$  and suppose  $A$  is  $\phi'$ -impredictable. We need to show  $\phi$ -impredictability where  $\phi(n) \leq \phi'(n)$  for all large enough  $n$ . Clearly changing finitely many initial values does not change  $\phi$ -impredictability, so we may assume  $\phi \leq \phi'$ . Let  $\psi \in \mathcal{T}$  be given by  $\phi'$ -impredictability, so

$$\forall \chi \in \mathcal{P} : \exists^\infty p : \psi(p) \in A \iff \chi(p, A \upharpoonright \phi'(p)) \downarrow$$

In particular, we can restrict to  $\hat{\chi} \in \mathcal{P}$  which given  $(p, w)$  first cuts off all but at most the first  $\phi(p) \leq \phi'(p)$  symbols of  $w$ , and then apply  $\chi$ . This restriction is

$$\forall \hat{\chi} \in \mathcal{P} : \exists^\infty p : \psi(p) \in A \iff \chi(p, A \upharpoonright \phi(p)) \downarrow,$$

which is just the definition of  $\phi$ -impredictability.

**Theorem 5.** For every  $\phi \in \mathcal{T}$ , there exists a  $\phi$ -impredictable  $\Sigma_1^0$  set.

*Proof.* We construct an increasing total computable function  $\psi$ , and  $A$  will be contained in its image. We want to have

$$\psi(p) \in A \iff \chi(p, A \upharpoonright \phi(p)) \downarrow$$

infinitely many times for all  $\chi$ . The way we construct  $A$  is we go through  $n \in \mathbb{N}$  and set either  $n \notin A$ , or set up the rule  $n \in A \iff \xi_n \downarrow$  for some Turing machine  $\xi_n$  whose behavior we describe (informally). We refer to those  $n \in \mathbb{N}$  already considered as *determined*.

List all partial functions  $\chi$  in an infinite-to-one way, i.e. we consider functions  $\chi$  successively, so that each  $\chi$  appears infinitely many times. When considering  $\chi$ , we want to determine new values in  $A$  so as to make sure that

$$\psi(p) \in A \iff \chi(p, A \upharpoonright \phi(p)) \downarrow$$

for at least one new  $p \in \mathbb{N}$ . Suppose we have already determined the first  $m$  values of  $A$ , i.e. the word  $w \in \{0, 1\}^m$  such that the characteristic sequence

of  $A$  will begin with the word  $w$  is already determined, and that we have not determined whether  $n \in A$  for any  $n \geq m$ . The idea is that while we do not know what the word  $w$  actually is, there are only  $2^m$  possible choices, and we simply try all of them to get the equivalence above to hold for one new  $p$ . To achieve this we will determine whether  $n \in A$  for some interval of choices  $n \in [m, M]$ .

Enumerate the words of  $\{0, 1\}^m$  as  $w_0, w_1, \dots, w_{2^m-1}$ . Now, let  $p_0 < p_1 < \dots < p_{2^m-1}$  be minimal such that  $\phi(p_i) \geq m$  for all  $i$ . Let  $M = \max_j \phi(p_j)$ . Set  $n \notin A$  for all  $n \in [m, M]$ . As  $\psi(p_i)$  for  $i \in [0, 2^m - 1]$  pick any distinct values greater than  $M$ , and determine  $\psi(p_i) \in A \iff \chi(p_i, w_i \cdot 0^{\phi(p_i)-m}) \downarrow$ . If  $w = w_i$ , then for  $p = p_i$  we have

$$\psi(p) \in A \iff \chi(p_i, w_i \cdot 0^{\phi(p_i)-m}) \downarrow \iff \chi(p, A \upharpoonright \phi(p)) \downarrow,$$

because the characteristic sequence of  $A$  indeed begins with the word  $w_i \cdot 0^{\phi(p_i)-m}$ . Thus, we have obtained a new value  $p = p_i$  at which the statement is satisfied for  $\chi$ .

For all  $n$  that are yet undetermined, but some larger number is determined, we determine  $n \notin A$ , so that we determine the values in some new interval  $[m, M']$ . We can then inductively continue to the next value of  $\chi$ . This process determines all values of  $A$ , and by construction  $A$  is a recursively enumerable set which is  $\phi$ -impredictable.

**Lemma 4.** *Let  $A, B \subset \mathbb{N}$  and suppose  $A$  many-one reduces to  $B$ , and  $B$  wtt-reduces to  $A$  with rate  $\beta$ . If  $A$  is  $\phi$ -impredictable, then  $B$  is  $(\beta \circ \phi)$ -impredictable.*

*Proof.* Let  $\phi' = \beta \circ \phi$ . Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be the many-one reduction from  $A$  to  $B$ , and  $g : \{0, 1\}^* \rightarrow \{0, 1\}^*$  the wtt-reduction from  $B$  to  $A$  with rate  $\beta$ . Since  $A$  is  $\phi$ -impredictable there exists  $\psi : \mathbb{N} \rightarrow \mathbb{N}$  such that

$$\forall \chi \in \mathcal{P} : \exists^\infty p : \psi(p) \in A \iff \chi(p, A \upharpoonright \phi(p)) \downarrow.$$

Setting  $\psi' = f \circ \psi$ , we have, by the definition of  $f$ , that

$$\forall \chi \in \mathcal{P} : \exists^\infty p : \psi'(p) \in B \iff \chi(p, A \upharpoonright \phi(p)) \downarrow.$$

For  $\chi \in \mathcal{P}$  define  $\hat{\chi} \in \mathcal{P}$  as follows: Given  $(p, w)$ , compute  $g(w) = u$  and then evaluate  $\chi(p, u \upharpoonright \phi'(p))$ , so that

$$\hat{\chi}(p, A \upharpoonright \phi(p)) \downarrow \iff \chi(p, g(A \upharpoonright \phi(p)) \upharpoonright \phi'(p)) \downarrow,$$

where we observe that  $g(A \upharpoonright \phi(p)) \upharpoonright \phi'(p) = B \upharpoonright \phi'(p)$  (in particular  $g$  indeed halts with an output so the formula makes sense).

Specializing the first quantifier, we have

$$\forall \hat{\chi} \in \mathcal{P} : \exists^\infty p : \psi'(p) \in B \iff \hat{\chi}(p, A \upharpoonright \phi(p)) \downarrow,$$

equivalently

$$\forall \chi \in \mathcal{P} : \exists^\infty p : \psi'(p) \in B \iff \chi(p, B \upharpoonright \phi'(p)),$$

so  $\psi'$  proves that  $B$  is  $\phi'$ -impredictable, as desired.

We show that unpredictability implies uncomputability.

**Proposition 1.** *If  $A$  is  $\Pi_1^0$ , then  $A$  is not  $\phi$ -uncomputable for any  $\phi \in \mathcal{T}$ .*

*Proof.* Suppose for a contradiction that  $A$  is  $\Pi_1^0$  and is  $\phi$ -uncomputable with  $\phi \in \mathcal{T}$ . Let  $\psi \in \mathcal{P}$  be such that

$$\forall \chi \in \mathcal{P} : \exists^\infty p : \psi(p) \in A \iff \chi(p, A \upharpoonright \phi(p)) \downarrow .$$

In particular, this applies to the following  $\chi$ : given  $(p, w)$ , we ignore  $w$  and if  $\psi(p) \notin A$ , then  $\chi(p, w) \downarrow$ , and otherwise  $\chi(p, w) \uparrow$ . This well-defines  $\chi \in \mathcal{P}$  since  $A$  is  $\Pi_1^0$ , and for all  $p \in \mathbb{N}$  we have

$$\psi(p) \in A \iff \chi(p, A \upharpoonright \phi(p)) \uparrow ,$$

a contradiction.

## 4 Impredictable torsion groups

**Definition 2.** *We define a group  $K(G, A, H)$  which depends on a choice of a finitely-generated groups  $G, H$  (and choices of generating sets for them, kept implicit), and a set  $A \subset \mathbb{N}$ . First define a subshift on  $G$  by*

$$X_A = \{x \in \{0, 1\}^G \mid \sum_{g \in G} x_g \leq 2, \text{ and } \sum_{k \in \{g, h\}} x_k = 2 \implies d(g, h) \notin A\}$$

where  $d : G \times G \rightarrow \mathbb{N}$  is the (left-invariant) word metric. To each  $g \in G$  associate the bijection

$$\hat{g}(x, h') = (g \cdot x, h'), \quad \hat{g} : X_A \times H \rightarrow X_A \times H$$

i.e. the usual shift action in the first component, and to each  $b \in \{0, 1\}$  and  $h \in H$  associate the bijection

$$h_b(x, h') = (x, h^{1-|x_{e_G}-b|} \cdot h'), \quad h_b : X_A \times H \rightarrow X_A \times H$$

where  $h^{1-|x_{e_G}-b|}$  evaluates to  $e_H$  if  $x_{e_G} = b$ , and to  $h$  otherwise. Define

$$K(G, A, H) = \langle \{\hat{g}, h_b \mid g \in G, h \in H, b \in \{0, 1\}\} \rangle \leq \text{Sym}(X_A \times H)$$

Obviously  $K = K(G, A, H)$  is finitely-generated, and the implicit generators we use for  $K$  are the ones from the definition, with  $\hat{g}$  taken only for  $g$  in the generating set of  $G$ , and  $h_b$  only for  $h$  in the generating set of  $H$ .

*Remark 1.* This group can be interpreted as a group of finite-state machines in the sense of [3] (with obvious nonabelian generalization) when  $H$  is finite, by simulating the actions of  $\hat{g}$  by translations of the head, having  $|H|$  states, and changing the state by the left-regular action of  $H$  (if in the correct clopen set) when  $h_b$  is applied. Again if  $H$  is finite,  $K$  can also be interpreted as a subgroup of the topological full group of the  $G \times H$ -subshift  $X_A \times H$  under the action  $(g, h) \cdot (x, h') = (gx, hh')$ , by having  $\hat{g}$  act by  $(g, e_H)$  and having  $h_B$  act by either  $(e_G, h)$  or  $(e_G, e_H)$  depending on  $x_{e_G}$ .

We now prove some important technical properties of these groups, leading up to the proof that they give examples of bounded torsion groups with unpredictable word problem.

**Lemma 5.** *If  $A' \subset A$ , the defining action of  $K(G, A, H)$  can be seen as a restriction of the action of  $K(G, A', H)$  in a natural way, thus  $K(G, A, H)$  is a quotient group of  $K(G, A', H)$ .*

*Proof.* Map generators to generators in the obvious way. The group  $K(G, A', H)$  acts on  $X_{A'} \times H \supset X_A \times H$  and the restriction of the action to  $X_A \times H$  is precisely that of  $K(G, A, H)$ . Thus, all identities of  $K(G, A', H)$  are identities also in  $K(G, A, H)$ .

In particular, the defining action of  $K(G, A, H)$  is always a restriction of  $K(G, \emptyset, H)$ .

**Lemma 6.** *For any  $G, A, H$ , there exists a split epimorphism  $K(G, A, H) \rightarrow G$ .*

*Proof.* We have  $X_{\leq 1} = X_{\mathbb{N}} \subset X_A$ , so the  $\hat{g}$ -translations act nontrivially on the  $X_A$  component as the action on  $X_{\leq 1}$  is the one-point compactification of the left-regular action of  $G$  on itself. Observe also that the  $X_A$ -component is not modified by any of the maps  $h_b$ , so the action of  $K_A$  on this component factors is just the shift action  $G \curvearrowright X_A$ . This gives a homomorphism  $\gamma : K(G, A, H) \rightarrow G$ , and the map  $g \mapsto \hat{g}$  is a section for it.

We refer to  $\gamma$  as the *natural epimorphism*.

**Lemma 7.** *If  $\text{WP}(G)$  is decidable and  $A$  is  $\Sigma_1^0$ , then  $X_A$  is an effective subshift.*

*Proof.* Since we can list elements of  $A$ , and can compute the distance between given group elements, we can forbid all finite patterns where two 1s appear at  $g, h$  with  $d(g, h) \in A$ .

**Lemma 8.** *If  $G$  has decidable word problem and  $H$  is recursively presented, then the following statements hold.*

- If  $H$  has decidable word problem, then the word problem of  $K(G, A, H)$  conjunctively reduces to  $A$  with exponential rate.
- If  $A$  is  $\Sigma_1^0$  then  $K(G, A, H)$  is recursively presented.

*Proof.* Let  $K = K(G, A, H)$  and let  $S$  be the finite generating set of  $K$ . We begin with the proof of the latter item. We need to find a semi-algorithm that, given  $w \in S^*$ , halts if and only if  $w$  represents the identity. For this, first consider the natural epimorphism image  $\gamma(w) \in G$ . Because  $G$  is in particular recursively presented, we can first verify that  $\gamma(w) = e_G$  (if not, then also  $w \neq e_K$ , and the computation diverges as desired).

Assuming  $\gamma(w) = e_G$ , we next check that  $w$  acts trivially on all elements of  $X_A \times H$ . We define an action of the free group on generators  $S$  on pairs  $(P, h) \in \mathcal{L}(X_\emptyset) \times H$  by

$$\hat{g}(P, h') = (g \cdot P, h'),$$

(where  $\text{dom}(gP) = g\text{dom}(P)$  and  $gP_k = P_{g^{-1}k}$ ), and for  $b \in \{0, 1\}$  and  $h \in H$  we map

$$h_b(P, h') = (P, h^{1-|P_{e_G}-b|}h'),$$

when  $e_G \in \text{dom}(P)$ , and  $h_b(P, h') = (P, h')$  otherwise.

If  $|w| \leq n$  and  $\gamma(w) = e_G$ , clearly  $w \approx e_K$  if and only if the action of  $w$  fixes all  $P \times h$  where  $P \in \mathcal{L}_n(X_A)$ ,  $h \in H$ . Naturally, if  $\mathcal{P} \supset \mathcal{L}_n(X_A)$  and the action fixes  $(P, h)$  for all  $P \in \mathcal{P}$ ,  $h \in H$ , then a fortiori  $w \approx e_K$ . We can verify this for a particular  $(P, h)$  by using the fact  $H$  is recursively presented. By the previous lemma,  $X_A$  is effective, so we can enumerate upper approximations to  $\mathcal{L}_n(X_A)$  which eventually converge. In other words, we eventually obtain the set  $\mathcal{L}_n(X_A)$ , and if  $w \approx e_K$ , then at this point (at the latest) we can conclude that  $w$  indeed acts trivially and halts.

The proof of the first item is similar. To see that there is a wtt-reduction with exponential rate, observe that if we know the first  $n$  values of  $A$ , then we can determine the legal contents of all  $G$ -patterns with domain  $B_n(G)$  in  $X_A$ , and using this, and the decidable word problem of  $H$ , we can determine whether  $w \approx e_K$  for any word with  $|w| \leq n$ . Since we list elements of groups in lexicographic order, the resulting rate  $\beta$  is exponential, as there are exponentially many words  $w$  with  $|w| \leq n$ .

To see that this is a conjunctive reduction, we observe that the reduction function  $g : \{0, 1\}^* \rightarrow \{0, 1\}^*$  (computing initial segments of the word problem from initial segments of  $A$ ) of the previous paragraph can be written uniformly for all sets  $A$  so it is total computable, and that we should set  $g(u)_i = 1$  if and only if the  $i$ th group element acts trivially on the set of patterns not containing elements of  $A$ . It can be checked by a terminating computation whether the group element acts nontrivially on some pattern containing only one 1. Our query on  $A$  should check that whenever the element acts nontrivially on a pattern with two 1s, then  $A$  contains the distance between the 1s of the pattern. This corresponds to checking  $u \geq w_{|u|, i}$  where  $w_{|u|, i}$  lists these finitely many bad distances.

**Lemma 9.** *Suppose  $G$  has decidable word problem and  $H$  is not abelian. Then  $A$  many-one reduces to the word problem of  $K(G, A, H)$ .*

*Proof.* For  $n \geq 1$ , let  $g_n$  be any effective list of elements of  $G$  satisfying  $|g_n| = n$  (using the fact  $G$  has decidable word problem), and consider  $g'_n = [h'_1, h_1^{g_n}]$  where  $[h', h] \neq e_H$  (using that  $H$  is not abelian) and the action of  $h_1^{g_n} = g_n h_1 g_n^{-1}$  is

$$h_1^{g_n}(x, h') = (x, h^{1-|x_{g_n}-b|} \cdot h')$$

We have  $g'_n \approx e_K$  if and only if  $n \in A$ . Namely, if  $n \notin A$  then  $g'_n$  acts nontrivially on  $(x, e_H)$  where  $x \in X_A$  is the unique configuration satisfying  $x_{e_G} = x_{g_n} = 1$ , while if  $n \in A$  then for all  $x \in X_A$  either  $x_{e_G} = 0$  or  $x_{g_n} = 0$ , and in either case a direct computation shows  $g'_n(x, h) = (x, h)$  for all  $h \in H$ .

**Lemma 10.** *For any torsion groups  $G, H$ , we have  $T_{K(G, A, H)} = O(T_G T_H)$ .*

*Proof.* Let  $K = K(G, A, H)$ . If  $w \in K$  with  $|w| \leq n$ , then  $\gamma(w) \in G$  with  $|\gamma(w)| = m \leq n$ . Let  $k = T_G(m)$  (note that  $k \mid T_G(n)$ ), so  $\gamma(w^k) = \gamma(w)^k = e_G$ . If  $w^k$  has order  $\ell$  for all  $w$ , then  $w^{\ell k} = (w^k)^\ell = e_K$ , and thus  $T_K(n) \leq \ell T_G(m)$ , and we have shown  $T_K = O(T_G T_H)$  as claimed.

So suppose that  $\gamma(w) = e_G$ , and consider the action on  $(x, h) \in X_A \times H$ . The action of  $w$  shifts  $x$  around, and based on its contents multiplies  $h$  from the right by elements of  $H$ . For any fixed  $x \in X_A$ ,  $\gamma(w) = e_G$  implies that there exists  $h_x \in H$  with  $|h_x| \leq |w|$ , such that  $w \cdot (x, h) = (x, h_x h)$  for all  $h \in H$ . Since  $h_x$  has order at most  $T_H(|h'|)$ , and  $T_H(|h'|) \mid T_H(|w|)$ , we have  $w^{T_H(|w|)} \cdot (x, h) = (x, h_x^{T_H(|w|)} h) = (x, h)$ , concluding the proof.

**Theorem 6.** *Let  $A \subset \mathbb{N}$  be  $\Sigma_1^0$ . For any f.g. torsion group  $G$  with decidable word problem there exists a recursively presented torsion group  $K$  with  $T_K = \Theta(T_G)$ , such that the word problem of  $K$  conjunctively reduces to  $A$  with exponential rate and  $A$  many-one reduces to the word problem of  $K$ .*

*Proof.* By the previous lemmas, if  $K = K(G, A, H)$  for finitely-generated groups  $G$  and  $H$  which have decidable word problems, and  $H$  is nonabelian with bounded torsion, then

- $K$  is recursively presented (Lemma 8),
- $T_K \geq T_G$  because  $G \leq K$  and by the choice of generators (Lemma 6),
- $T_K \leq T_G T_H = O(T_G)$  (Lemma 10),
- $\text{WP}(K)$  conjunctively reduces to  $A$  with exponential rate (Lemma 8),
- $A$  many-one reduces to  $\text{WP}(K)$  (Lemma 9).

In the previous theorem, the implicit constants for  $\Theta$  can be taken to be 1 (for the lower bound) and 6 (for the upper bound, by setting  $H = S_3$ ). Of course, if  $G$  has bounded torsion, so does  $K$ .

**Theorem 7.** *Let  $\phi$  be a total recursive function. Then there exists a recursively presented torsion group with bounded torsion, whose word problem is  $\phi$ -impredictable.*

*Proof.* Let  $A$  be a  $\phi$ -impredictable  $\Sigma_1^0$  set (Theorem 5) and apply the previous theorem to obtain a recursively presented torsion group  $K$  with bounded torsion, such that the word problem of  $K$  conjunctively reduces to  $A$  with exponential rate  $\beta$  and  $A$  many-one reduces to the word problem of  $K$ . Then in particular the word problem of  $K$  wtt-reduces to  $A$  with exponential rate  $\beta$ . By Lemma 4, the word problem of  $K$  is  $(\beta \circ \phi)$ -impredictable. In particular, it is  $\phi$ -impredictable by Lemma 3.

## 5 Application: Four heads are better than three

We first define group-walking automata and the subshifts they recognize. By  $\pi_i$  we mean the projection to the  $i$ th coordinate of a finite Cartesian product. For  $Q_i \not\ni 0$  a finite set write  $X_{Q_i}^1$  for the subshift on a group  $G$  clear from context containing those  $x \in (Q_i \cup \{0\})^G$  satisfying  $|\{g \in G \mid x_g \neq 0\}| \leq 1$ .

**Definition 3.** Let  $\Sigma$  be a finite alphabet. A  $k$ -headed group-walking automaton on the full shift  $\Sigma^G$  is a tuple  $\mathcal{A} = (\prod_{i=1}^k Q_i, f, I, F, S)$ , where  $Q_1, Q_2, \dots, Q_k$  are state sets not containing the symbol 0,  $I$  and  $F$  are finite clopen subsets of the product subshift  $Y = \prod_{i=1}^k X_{Q_i}^1$ , and  $f : \Sigma^G \times Y \rightarrow \Sigma^G \times Y$  is a cellular automaton satisfying  $\pi_1 \circ f = \pi_1$  and

$$\pi_i(\pi_2(f(x, y))) = 0^G \iff \pi_i(y) = 0^G$$

for all  $x \in \Sigma^G$ ,  $y \in Y$  and  $i \in \{1, \dots, k\}$ .

For a  $k$ -headed automaton  $\mathcal{A}$  as above, we denote by  $\mathcal{S}(\mathcal{A}) \subset \Sigma^G$  the subshift

$$\{x \in \Sigma^G \mid \forall g, h \in G, y \in I, n \in \mathbb{N} : h \cdot \pi_2(f^n(g \cdot x, y)) \notin F\}.$$

For  $k \geq 1$ , we denote by  $\mathcal{S}(G, k)$  the class of all subshifts  $\mathcal{S}(\mathcal{A})$  for  $k$ -headed automata  $\mathcal{A}$ , and  $\mathcal{S}(G, 0)$  is the class of all  $G$ -SFTs. We also write  $\mathcal{S}(G) = \bigcup_{k \in \mathbb{N}} \mathcal{S}(G, k)$ .

This definition may seem cryptic on a first reading. Its details are unraveled in [15] (see [14] for a discussion of possible variants). Our interpretation of  $X \in \mathcal{S}(G, k)$  is that a  $k$ -headed group-walking automaton can define  $X$ , for a particular (in our opinion natural) way of defining subshifts by such automata.

Key points are that interpreting the  $i$ th track of  $Y$  as giving the position or a head and its current state,  $f$  is a local rule that tells how the heads move on configurations, and the assumptions imply that all heads are always present, are initialized in (roughly) the same position (described by  $I$ ), have to join together to reject a configuration (described by  $F$ ), and cannot communicate over distances (because  $f$  is a cellular automaton).

The following observation is essentially Proposition 2 in [15], though here we “complement” the separating subshift, because the result of [15] cannot be used with recursively presented groups.

**Lemma 11.** Let  $G$  be a finitely-generated torsion group, let  $T(n) = \max(n, T_G(n))$ , and suppose that there exists a superexponential function  $\zeta : \mathbb{N} \rightarrow \mathbb{N}$  such that the word problem of  $G$  is  $(\zeta \circ \zeta \circ T \circ \zeta \circ T \circ \zeta)$ -impredictable. Then

$$\mathcal{S}(G \times \mathbb{Z}, 3) \subsetneq \mathcal{S}(G \times \mathbb{Z}, 4).$$

We sketch the proof from [15], for our complemented definitions.

*Proof (Proof sketch).* Let  $\zeta$  be superexponential and recursive, let  $G$  be such that the word problem of  $G$  is  $\phi$ -impredictable for  $\phi = \zeta \circ \zeta \circ T \circ \zeta \circ T \circ \zeta$ , and let  $\psi$  be the corresponding function, so

$$\forall \chi \in \mathcal{P} : \exists^\infty p : \psi(p) \sim e_G \iff \chi(p, G \upharpoonright \phi(p)) \downarrow.$$

For each  $p \in \mathbb{N}$ , let  $x^p \in \{0, 1\}^{G \times \mathbb{Z}}$  be the configuration where  $x_{(g,n)}^p = 1$  if and only if  $n \equiv 0 \pmod{p}$ . Define  $B = \{p \in \mathbb{N} \mid \psi(p) \sim e_G\}$ , and let  $Y_B \subset \{0, 1\}^{G \times \mathbb{Z}}$  be the smallest subshift containing the configurations  $x^p$  with

$p \in B$ , i.e. the forbidden patterns are (encodings of) the complement of the word problem of  $G$  (together with the recursive set of patterns ensuring  $Y_B \subset Y_\emptyset$ ). We clearly have  $x^p \in Y_B \iff p \in B \iff \psi(p) \sim e_G$ .

The crucial observation in [15] was that for any fixed three-headed automaton  $\mathcal{A}$ , the function  $(\zeta \circ T \circ \zeta \circ T \circ \zeta)(p)$  eventually bounds how far the  $G$ -projections of the heads can be from each other during valid runs on a configuration  $x^p$ , which means that we can construct a Turing machine that, given access to an oracle for the initial segment  $G \upharpoonright \phi(p)$  of the word problem (we add an extra  $\zeta$  since we linearize the word problem), we can simulate all runs of  $\mathcal{A}$  on the configurations  $x^p$  (observe that there are essentially only  $p$  different starting positions that need to be considered), halting if and only if one of them halts (i.e. the finite clopen set  $F$  is entered and the configuration is forbidden).

Suppose for a contradiction that  $\mathcal{A}$  is a three-headed automaton that defines  $Y_B$ . Letting  $\chi$  be the Turing machine described above which simulates  $\mathcal{A}$ , we have for all large enough  $p$  that

$$\chi(p, G \upharpoonright \phi(p)) \downarrow \iff x^p \notin Y_B$$

but by the definition of  $\psi$  there exist arbitrarily large  $p$  such that

$$\psi(p) \sim e_G \iff \chi(p, G \upharpoonright \phi(p)) \downarrow \iff x^p \notin Y_B \iff \psi(p) \not\sim e_G,$$

a contradiction.

On the other hand  $Y_B \in \mathcal{S}(G \times \mathbb{Z}, 4)$  since it is intrinsically  $\Pi_1^0$  in the sense of [15], by a similar proof as in [15], observing that using an oracle for the word problem of  $G$  we can easily forbid all  $x^p$  with  $p \notin B$ .

Four heads are now seen to be better than three:

*Proof (Proof of Theorem 3).* By Theorem 7, there exists a group  $G$  which has bounded torsion and has  $\exp \circ \exp \circ \exp \circ \exp \circ \exp$ -impredictable word problem. Such a group is  $(\zeta \circ \zeta \circ T \circ \zeta \circ T \circ \zeta)$ -impredictable for some superexponential function  $\zeta$ , where  $T(n) = n$ , and thus the previous lemma implies  $\mathcal{S}(G \times \mathbb{Z}, 3) \subsetneq \mathcal{S}(G \times \mathbb{Z}, 4)$ .

*Remark 2.* For maximal ‘‘automaticity’’,<sup>1</sup> or to avoid using f.g. bounded torsion groups (whose infiniteness is rather difficult to verify for mortals), one can replace the use of bounded torsion groups by any torsion group with recursive torsion function. For example using the Grigorchuk group as  $G$  in the construction of  $K$  (then  $T_K(n) \leq O(n^3)$  by [5, Theorem VIII.70]), and using Theorem 6 directly, the previous proof goes through with the same exponential tower. Automata groups in the sense of [4] cannot have bounded torsion by Zelmanov’s theorem [5, 16, 17].

---

<sup>1</sup> Using the Grigorchuk group or another torsion automata group, the statement of our main result is somewhat amusing, in that it states that a group of finite-state **automata** acting on a subshift on an **automata** group admits subshifts definable by group-walking **automata** with four heads but not three.

## References

1. Sergei I. Adian. The Burnside problem and related topics. *Russian Mathematical Surveys*, 65(5):805–855, jan 2011.
2. Nathalie Aubrun, Sebastián Barbieri, and Mathieu Sablik. A notion of effectiveness for subshifts on finitely generated groups. *Theoretical Computer Science*, 661(Supplement C):35 – 55, 2017.
3. Sebastián Barbieri, Jarkko Kari, and Ville Salo. *The Group of Reversible Turing Machines*, pages 49–62. Springer International Publishing, Cham, 2016.
4. Laurent Bartholdi and Pedro V Silva. Groups defined by automata. *arXiv preprint arXiv:1012.1531*, 2010.
5. Pierre de La Harpe. *Topics in geometric group theory*. University of Chicago Press, 2000.
6. Rostislav Ivanovich Grigorchuk and Konstantin Medynets. On algebraic properties of topological full groups. *Sbornik: Mathematics*, 205(6):843–861, 2014.
7. Pierre Guillon, Emmanuel Jeandel, Jarkko Kari, and Pascal Vanier. Undecidable word problem in subshift automorphism groups. *CoRR*, abs/1808.09194, 2018.
8. Sergei V. Ivanov. The free Burnside groups of sufficiently large exponents. *International Journal of Algebra and Computation*, 4(01n02):1–308, 1994.
9. K. Juschenko and N. Monod. Cantor systems, piecewise translations and simple amenable groups. *ArXiv e-prints*, April 2012.
10. Hiroki Matui. Some remarks on topological full groups of cantor minimal systems. *International Journal of Mathematics*, 17(02):231–251, 2006.
11. Alexei Myasnikov and Denis Osin. Algorithmically finite groups. *Journal of Pure and Applied Algebra*, 215(11):2789 – 2796, 2011.
12. Ville Salo. Impredictable subsets of  $\mathbb{N}$ . MathOverflow. <https://mathoverflow.net/questions/354526/> (version: 2020-3-11).
13. Ville Salo and Michael Schraudner. Automorphism groups of subshifts through group extensions. Preprint.
14. Ville Salo and Ilkka Törmä. Plane-walking automata. In Teijiro Isokawa, Katsunobu Imai, Nobuyuki Matsui, Ferdinand Peper, and Hiroshi Umeo, editors, *Cellular Automata and Discrete Complex Systems*, volume 8996 of *Lecture Notes in Computer Science*, pages 135–148. Springer International Publishing, 2015.
15. Ville Salo and Ilkka Törmä. Independent finite automata on Cayley graphs. *Natural Computing*, 16(3):411–426, Sep 2017.
16. Efim Isaakovich Zel'manov. A solution of the restricted Burnside problem for 2-groups. *Matematicheskii Sbornik*, 182(4):568–592, 1991.
17. Efim Isaakovich Zel'manov. Solution of the restricted Burnside problem for groups of odd exponent. *Mathematics of the USSR-Izvestiya*, 36(1):41, 1991.