ALGORITHMIC PROBLEMS IN GROUPS, SEMIGROUPS AND INVERSE SEMIGROUPS

S. MARGOLIS, J.MEAKIN, M.SAPIR
*University of Nebraska at Lincoln**

Contents

* This research was supported in part by the NSF grant DMS-9203981 and by the Center for Communication and Information Science, University of Nebraska-Lincoln

1. Introduction

Algorithmic problems in modern algebra had their origins in logic and topology - in the works of Thue [123], Tietze [120] and Dehn [25] at the beginning of the century. They showed that the problem of deducibility of relations in associative calculi, the homeomorphism problem for topological manifolds and the homotopy equivalence problem in finite dimensional manifolds all turned out to be equivalent to algebraic problems, namely the word problem for finitely presented semigroups and groups and the isomorphism and conjugacy problems for finitely presented groups.

Let $\mathcal{V}$ be a variety of universal algebras, let F_X be the finitely generated free object in $\mathcal{V}$ with a set X of free generators. Let σ be a congruence on F_X generated by finitely many pairs $R = \{(u_i, v_i) : i = 1, 2, \dots\}$. Then the factor algebra F_X/σ is called *finitely presented inside* $\mathcal{V}$. We will denote this algebra by $< X : R, \mathcal{V} >$. If it is clear what variety $\mathcal{V}$ we are dealing with, we will omit $\mathcal{V}$ from this notation. We say that the word problem is *decidable (solvable)* in F_X/σ if there exists an algorithm which tells us for every pair of elements (x, y) from F_X if $(x, y) \in \sigma$.

As is well known the unsolvability of the word problem for semigroups was proved in 1947 by A.A.Markov [61] and E.L.Post [91] and significantly more difficult results on the unsolvability of the three group problems: the word problem, the conjugacy problem and the isomorphism problem — were proved by P.S.Novikov [78] at the beginning of the fifties. A few years later W.W.Boone [16] gave another proof of the unsolvability of the word problem in groups.

The investigation of algorithmic problems, stimulated initially from problems in logic and topology, is largely motivated now by the internal needs of algebra. Algorithmic problems lie very often at the heart of difficult algebraic problems. For example, central to the deep results on Burnside problems for periodic groups is the solution to the word and conjugacy problems in the Burnside groups [79], [2], [82], [42], [51].

While the main algorithmic problems were proved to be unsolvable in general, further investigations have been carried out in two directions. The first direction deals with algebras given by defining relations which satisfy certain restrictions, for example groups and semigroups with small cancellation restrictions, hyperbolic groups, automatic groups and semigroups, groups and semigroups given by a small number of relations. There are many deep theorems obtained in this direction (see [50], [34], [29]) and many deep problems. Among these problems one can mention the word problem for one-relator semigroups. It is known that every semigroup given by a relation of the form $w = 1$ has solvable word problem, but the general case is still unsolved in spite of 30 years of continuous attempts. For other partial results see the surveys [3] and [62]. Recall that in the case of groups Magnus' theorem states that the word problem is solvable in every one-relator group [50].

Another direction investigates algorithmic problems in some subclasses of semigroups, groups, etc. Algorithmic problems are important from a general, "philosophical", point of view also. Thus, the solvability of the word problem in a class of algebraic systems usually means that it is possible to study structural properties of systems in this class. Conversely unsolvability of the word problem usually means that there will be major difficulties in the investigation of this class "as a whole". Among all possible subclasses, the most natural and important ones are, of course, varieties – classes given by identities. The remarkable role they play in algebra and the fact that it is very natural to consider algorithmic problems in varieties induced a large and robust interest in this branch of algebra during the last 15-20 years. This study occupied a major part in most surveys devoted to algorithmic problems in algebra (see e.g. [15], [3], [95], [14], [45]).

Varieties are the most intensively studied classes of universal algebras because they combine the most natural operations on both syntactic and semantic levels. Birkhoff's fundamental theorem connecting these two points of view leads to a rich interplay between combinatorics on words and structural methods. We recall the basic definitions here.

Syntactic: a variety of algebras of a given type is a collection of algebras all satisfying a given set of identities.

Semantic: a variety of algebras of a given type is a collection of algebras closed under the operations of taking subalgebras, quotients and direct products.

Birkhoff's theorem then says that these two definitions define the same classes of algebras. In this paper we will be interested in finitely described varieties of groups, semigroups and inverse semigroups. Clearly, there is both a syntactic and semantic notion of the term "finitely described".

Syntactically finitely described: a variety is finitely based if it is defined by a finite set of identities.

Semantically finitely described: a variety is finitely generated if it is the smallest variety containing a fixed finite algebra.

Given a property α of varieties one can ask for an algorithm which decides if a finitely described variety has the property α. If we find such an algorithm then we will say that we have an *algorithmic description* of those varieties that satisfy our property α.

The last ten years have yielded algorithmic descriptions of varieties with many important properties. Among all properties α of semigroup varieties which have been studied, two turned out to be of major importance: the Burnside property and the finite basis property (see definitions later in the paper). In this paper we will show surprisingly close connections between these properties and other properties.

During the last 20-30 years another source of algorithmic problems appeared, namely theoretical computer science. In particular the theory of finite automata and formal languages has served as a source of many algorithmic problems. Several algorithmic problems about the formal language aspects of finite automata are decidable, while their computational complexity tends to be high. For example, the question as to whether two non-deterministic finite automata recognize the same language is PSPACE complete. Many other problems, especially those concerned with the semigroup theoretic aspects of the theory of finite automata, are undecidable. Further influences included an analysis of the complexity of algorithms in the case where the algorithmic problem is decidable. For example it is known that the word problem in any commutative semigroup may be solved in polynomial time. There is also a very large literature devoted to efficient algorithms for computations in finite groups (see [4]) and other areas of algebra.

The present paper is concerned primarily with a study of algorithmic problems in semigroups, groups and inverse semigroups. The study of such problems for semigroups and groups is a classical part of algebra. Inverse semigroups were introduced into the literature in the papers [124] and [92] in the early 1950's. They may be regarded as semigroups of partial one - one transformations (see [88]). Algorithmic problems for inverse semigroups are natural and have received considerable attention in the literature during the past 15 - 20 years. While algorithmic problems for semigroups and groups are basically problems about words, the corresponding problems for inverse semigroups are essentially problems about finite labelled trees. Indeed the free inverse semigroup may be regarded as a semigroup of finite (birooted) labelled trees (see [75]). There is a developing theory of varieties of inverse semigroups (see [88] for a brief introduction) and of algorithmic problems for finitely presented inverse semigroups. Inverse semigroups form in a sense an intermediate class between groups and semigroups and have influenced both group theory and semigroup theory. For example, congruences on free monoids may be classified via inverse submonoids of the polycyclic monoid [67], and finitely generated subgroups of free groups may be classified via finite inverse semigroups [56], [12]. We will provide some details of these and other connections between inverse semigroups, groups and semigroups later in the paper.

We shall employ geometric, topological, and combinatorial methods to study algorithmic problems for groups, semigroups, and inverse semigroups.

2. Some Geometric Methods

2.1. Groups

Geometric methods have played an important role in combinatorial group theory since the inception of work in that field. We shall not attempt to provide a detailed survey of such methods here, but rather we draw the reader's attention to two standard general ways in which geometric methods have been used to study algorithmic problems for group presentations.

Let $G = gp < X : R >$ be a presentation of a group G. Here X is a non-empty set and we may assume that $R = \{R_i : i \in I\}$ is a (possibly empty) set of non-trivial cyclically reduced words in $(X \cup X^{-1})^*$. Thus $G = FG(X)/N$, where $FG(X)$ denotes the free group on X and N is the normal closure of R in $FG(X)$. Denote by π the canonical projection of $FG(X)$ onto G. We shall adhere to this notation throughout this section.

There are various ways to construct 2-complexes from the presentation. For example we denote by $K = K(X : R)$ the two-dimensional CW-complex with a single 0-cell whose 1-cells are in one-one correspondence with X, and whose 2-cells are in one-one correspondence with R, a given 2-cell being attached by the boundary path determined by reading the corresponding member of R. The underlying 1-skeleton (graph) of K is B_X, the bouquet of X circles. It is well known that $\pi_1(K) = G$ (see [50] or [23]).

It is also standard to consider the universal cover $\tilde{K} = \tilde{K}(X : R)$ of K. This is constructed in the following way. We may take $G = \pi_1(K)$ as the set of 0-cells of $\tilde{K}$. Next, for each $g \in G$ and $x \in X$, we attach a 1-cell (g, x) joining g and $g\pi(x)$: we orient (g, x) from g to $g\pi(x)$ and write $(g\pi(x), x^{-1})$ for the corresponding inverse. Then for each $g \in G$ and $r \in R$ we attach a 2-cell (g, r) via the boundary path $(g_1, x_1)(g_2, x_2)...(g_n, x_n)$ where $r = x_1x_2...x_n, x_i \in X \cup X^{-1}, g_1 = g, g_{i+1} = g_i\pi(x_i), i = 1, \ldots, n-1$. The 1-skeleton (graph) of $\tilde{K}$ is called the Cayley graph of the presentation and is denoted by $\Gamma(X : R)$. One obtains the Cayley complex $C(X : R)$ from $\tilde{K}(X : R)$ by identifying all the faces $(g, r), (g\pi(s), r), \ldots, (g\pi(s^{m-1}), r)$ where $r = s^m, m > 1$ and s is not a proper power in $FG(X)$. The 1-skeleton of $C(X : R)$ is again the Cayley graph $\Gamma(X : R)$ of course.

The 2-complexes described above (and in particular the Cayley complex and its 1-skeleton) play a prominent role in the theory. We refer the reader to the third chapter of [50] for basic information along these lines. In particular, the structure of groups with a presentation whose Cayley complex is planar is studied in [50] and it is shown that the word problem for such a presentation is decidable. Also in [50] the notion of asphericity of group presentations is introduced. In fact there are several somewhat different notions of asphericity that have been considered in the literature. We refer the reader to [19] where five notions of asphericity (all related to the complexes discussed above) are studied in detail.

It is easy to see that the word problem for the presentation $G = gp < X : R >$ is decidable if the ball B_n of radius n in the Cayley graph $\Gamma(X : R)$ is, for all n, finite and effectively constructable. There is a standard procedure (essentially the Todd-Coxeter method of coset enumeration) for inductively constructing $\Gamma(X : R)$ as a limit of a sequence of graphs Γ_i under graph morphisms $\phi_{i,j} : \Gamma_i \to \Gamma_j$ for all $i < j$. Since this procedure is closely related to much of what follows, we sketch a version of it here and at the same time introduce some notation and terminology that will be used in subsequent parts of the paper. The version that we present is a slight variation on the version discussed in [29]. There seem to be several versions of this algorithm that appear throughout the literature.

We restrict attention to the case that the presentation $G = gp < X : R >$ is finite (i.e X and R are finite). The idea is to build the Cayley graph iteratively from the origin (the vertex representing the identity of G) by using the relators to combine vertices and create loops. At each stage of the procedure we will have constructed a finite graph Γ_i and a graph morphism $\phi_i : \Gamma_i \to \Gamma_{i+1}$. The graphs Γ_i enjoy the following properties:

(P1) each graph Γ_i is connected with directed edges labelled by the elements of $X \cup X^{-1}$ and such that if there is an edge labelled by x from v_1 to v_2 , then there is also an edge labelled by x^{-1} from v_2 to v_1;

(P2) no two directed edges with the same initial vertex have the same label.

Graphs with these properties are referred to as "unambiguous partial Cayley graphs" in [29] and "inverse word graphs over X" in [118]. There is an evident notion of morphism between inverse word graphs (a graph homomorphism that preserves labels and directions). If Γ is an inverse word graph over X there is an obvious morphism $f : \Gamma \to B_X$ from Γ to the bouquet of X circles : f takes an edge labelled by x in Γ onto the edge labelled by x in B_X. It is clear that this morphism is locally injective at each vertex of Γ ; i.e f is an "immersion" in the sense of [116]. We sometimes abuse notation slightly and refer to Γ itself as an immersion over B_X. Graph immersions have been used by a number of authors ([116], [32], [12], [56]) to study subgroups of free groups and submonoids of free inverse monoids. We shall consider these ideas in more detail in the next section of this paper.

The reason for the name "inverse word graph over X" to describe a graph satisfying properties (P1) and (P2) above is that the transition monoid of such a graph (i.e. the monoid of partial transformations of the vertices of Γ induced by the action of the letters in $X \cup X^{-1}$) is an inverse monoid. An inverse monoid (semigroup) is a monoid (semigroup) M such that for each $x \in M$ there is a unique element (denoted by x^{-1}) in M such that

$$x = xx^{-1}x \text{ and } x^{-1} = x^{-1}xx^{-1}.$$

Such monoids may be viewed as monoids of injective (partial) functions on a set A. Indeed the monoid $SIM(A)$ of all injective partial functions on A (with respect to the usual composition of partial functions) is an inverse monoid (called the "symmetric inverse monoid on A") and the Preston-Wagner theorem asserts that every inverse monoid may be embedded in a suitable symmetric inverse monoid. We refer the reader to [88] for much basic information about inverse monoids. We shall return to a discussion of inverse monoids later in this paper.

We also mention at this stage that an inverse word graph Γ over X may be viewed as an automaton over the alphabet $X \cup X^{-1}$ if we distinguish an initial vertex (state) α and a terminal vertex (state) ω of Γ. Such an automaton is the minimal automaton of the language that it accepts and is referred to in this paper and in the literature as an "inverse automaton over X". The transition monoid of such an automaton is an inverse monoid (which is finite if Γ is finite). The inverse word graphs Γ_i that we are about to construct will have a fixed base point, which may be viewed as the initial and terminal vertex of Γ_i, so we may think of each Γ_i as an inverse automaton over X. Recall that an inverse automaton over an alphabet X is an automaton over $X \cup X^{-1}$ such that each $x \in X \cup X^{-1}$ induces an injective function on the state set and such that the inverse letter induces the inverse injective function.

In order to describe the graphs Γ_i associated with the finite presentation $G =$

$gp < X : R >$ we consider three types of operations that we will perform on a finite labelled graph Γ that satisfies property (P1):

(O1) Loop closings: if in Γ there is a path p from vertex v_1 to vertex $v_2 \neq v_1$ labelled by a relator $r \in R$, then we form a new graph Γ' obtained from Γ by identifying v_1 and v_2. Since R and Γ are finite and Γ' has one less vertex than Γ it follows that, by starting with Γ and applying successive loop closings a finite number of times, we will reach a (unique) graph Γ_l to which no further loop closings are applicable.

(O2) Adding hairs: if in Γ there is a vertex v and a letter $x \in X \cup X^{-1}$ such that Γ has no edge with label x starting at v, then we create a new graph Γ' from Γ by adding to Γ a new vertex v' and a new edge e with label x, initial vertex v and terminal vertex v' (and an inverse edge e' from v' to v with label x^{-1}). Since Γ and X are finite, there are only finitely many hairs that can be added to the vertices of Γ, thereby obtaining a new graph Γ_h.

(O3) Folding: if in Γ there are two (directed) edges with the same initial vertex (or terminal vertex) and the same label $x \in X \cup X^{-1}$, then we create a new graph Γ' from Γ by identifying these edges. One can check that folding is confluent, so that after finitely many foldings one reaches a graph Γ_f to which no further foldings apply: clearly Γ_f is an inverse word graph over X.

We construct a sequence of inverse word graphs Γ_0, $\Gamma_1, \ldots \Gamma_i$, associated with $G = gp < X : R >$ as follows. Start by taking a base point 1 and attaching to it a loop for each relator in R. Each loop starts and ends at 1 and successive edges describe the word in R. Next apply successive foldings until no more can be applied. Then close the resulting graph under loop closings. This may create more vertices at which foldings may be applied. Continue applying loop closings and foldings until no more loop closings or foldings may be applied. Finiteness of Γ, X and R and confluence of the operations guarantees that after finitely many steps we reach an inverse word graph Γ_0. If we regard the base point as the initial (terminal) vertex, then Γ_0 becomes an inverse automaton.

Assume inductively that we have constructed Γ_i. If, for each vertex v in Γ_i and each letter $x \in X \cup X^{-1}$, there is an edge in Γ_i with label x and initial vertex v, then G must be finite and Γ_i is the Cayley graph $\Gamma(X : R)$. Then define $\Gamma_{i+1} = \Gamma_i$ and ϕ_i to be the identity map on Γ_i. Otherwise we add all missing hairs to the vertices of Γ_i (obtaining the graph $(\Gamma_i)_h$) and then apply all possible loop closings and foldings until no more can be applied. Call the resulting graph Γ_{i+1}.

Note that each vertex and edge of Γ_i has a natural image in Γ_{i+1} after all the foldings and loop closings are applied, so there is a natural morphism $\phi_i : \Gamma_i \to \Gamma_{i+1}$. This gives rise to morphisms $\phi_{i,j} : \Gamma_i \to \Gamma_j$ for $i < j$ and it is not too difficult to see that $\Gamma(X : R)$ is the direct limit of this family $\{\Gamma_i, \phi_{i,j}\}$ of inverse word graphs and morphisms. If B_n is the ball of radius n in $\Gamma(X : R)$, then there is some integer i (depending on n, but in general not an effectively computable function of n) such that B_n is contained in Γ_i.

Theorem 2.1 *Let $G = gp < X : R >$ be a finitely presented group and let*

$\{\Gamma_i, \phi_{i,j}\}$ be the system of inverse word graphs and graph morphisms constructed according to the Todd-Coxeter algorithm described above. Then $\Gamma(X : R)$ is the direct limit of this system of graphs and graph morphisms.

The theorem essentially shows that every finitely presented group may be naturally approximated by finite inverse monoids (the transition monoids of the graphs Γ_i constructed above). In fact inverse monoids often arise naturally when given partial information about a group. We shall see other examples of this sort of phenomenon later in the paper.

We turn now to a very brief indication of a second standard way in which geometric methods have played a prominent role in combinatorial group theory. Again let $G = gp < X : R >$ be a finitely presented group, where the notation is as specified above. It is standard to study membership in N (the normal closure of R in $FG(X)$) by associating with each word in N a planar map (diagram) which is referred to in the literature as a van Kampen diagram (or a singular disk diagram). Briefly, the definition of such a diagram may be stated in the following way. We define a map to be a finite, planar, connected and simply connected 2-complex. A diagram Δ over an alphabet of the form $X \cup X^{-1}$ is a map such that every edge e (i.e., a 1-cell) of Δ is provided a label $\phi(e)$ in $X \cup X^{-1}$ such that $\phi(e^{-1}) = \phi(e)^{-1} \in (X \cup X^{-1})^*$. The label of a path $p = e_1 e_2 ... e_n$ in Δ is by definition the word $\phi(e_1)\phi(e_2)\dots\phi(e_n)$. We call a diagram Δ over $X \cup X^{-1}$ a (van Kampen) diagram over the presentation $G = gp < X : R >$ if the label of a boundary path of every face (i.e., a 2-cell) is a cyclic permutation of some relator r (or r^{-1}) in R. The van Kampen Lemma then says that a word $w \in (X \cup X^{-1})^*$ is in N if and only if there exists a diagram over G such that w is the label of the boundary of the diagram. In fact one may assume that the diagram is reduced. A diagram Δ is said to be reduced if there are no opposite faces in Δ. Two faces Π_1 and Π_2 are called "opposite" if there is an edge e in the intersection of their boundaries such that the labels of Π_1 and Π_2, starting at e, coincide. (That is, Π_1 and Π_2 are "mirror images" of each other). For a proof of this lemma, and for further definitions and results about van Kampen diagrams, we refer the reader to Chapter V of [50]. There are slightly different definitions that appear at various places in the literature. The reader is referred to Ol'shanskii's book [82] for a deep analysis of the use of these methods in combinatorial group theory. The methods of small cancellation theory [50] and much of the recent deep work of Ivanov on the Burnside problem [42] is based on these methods.

We remark that there is a natural graph morphism from any van Kampen diagram over G into the Cayley graph $\Gamma(X : R)$. We define such a map γ as follows. Let Δ be a van Kampen diagram over $G = gp < X : R >$ and suppose that p is a boundary path of Δ with initial (and terminal) vertex O. Define $\gamma(O) = 1$ (the vertex representing 1 in $\Gamma(X : R)$) and for an arbitrary vertex v of Δ let $\gamma(v)$ be the element of G represented by the word $\phi(q)$, where q is any path in Δ between the vertex O and v. In view of the van Kampen lemma the vertex $\gamma(v)$ of $\Gamma(X : R)$ depends only on v and not on the path q. For an edge e of Δ with initial vertex v let $\gamma(e)$ be the edge of $\Gamma(X : R)$ with initial vertex $\gamma(v)$ and label $\phi(e)$. It is clear that γ defines a graph morphism from (the 1-skeleton of) Δ into $\Gamma(X : R)$. Of course γ is not in general an embedding since distinct vertices (edges) of Δ may be identified under γ. The interplay between van Kampen diagrams and Cayley graphs indicated above is used throughout the literature in combinatorial group theory. For a nice recent application of these ideas we refer to Ol'shanskii's recent elementary proof [83] that groups with subquadratic isoperimetric inequality are hyperbolic.

2.2. Semigroups

We turn now to the use of geometric methods for semigroup presentations analogous to the methods discussed above for group presentations. Throughout this section we let $S = sgp < X : R >$ be a semigroup presentation. Here X is a non empty set and we will take R to be a finite set of relations of the form $u_i = v_i$ where u_i and v_i are non empty words in the free semigroup X^+. There is a well developed analogue of the planar van Kampen diagrams of combinatorial group theory in this setting and also an appropriate but somewhat less well developed analogue of the Cayley graph for a group presentation. We provide a brief discussion of both methods below. We begin with a description of the planar semigroup diagrams associated with the semigroup presentation S above.

A semigroup diagram (or s-diagram) over the presentation $S = sgp < X : R >$ for a pair (u, v) of non empty words in X^+ is a diagram Δ over $X \cup X^{-1}$ with the following properties:

(SD1) each face (i.e. bounded 2-cell) of Δ is labelled in the clockwise direction by a word of the form rs^{-1} for some (r, s) such that $r = s$ (or $s = r$) is a relation in R;

(SD2) the boundary of Δ carries a clockwise label uv^{-1};

(SD3) there are no interior sources or sinks (a source is a vertex of indegree 0 - i.e a vertex with no positively labelled edge coming in to it; a sink is the dual concept).

It is clear that if F is any face of such a diagram then there is a unique vertex F_- such that the clockwise boundary cycle of F starting at F_- is of the form $\alpha\beta^{-1}$, where $\phi(\alpha) = r$ and $\phi(\beta) = s$. We denote the terminal point of α (= the terminal point of β) by F_+ and refer to α [β] as the left [right] boundary of F. Similarly the whole diagram Δ has a left boundary labelled by u starting at Δ_- and a right boundary labelled by v ending at Δ_+. An important property of such diagrams is that every positively labelled edge of Δ lies on some positively labelled directed path running from Δ_- to Δ_+. (Such a path is sometimes referred to as a "transversal" of Δ.) The analogue of the van Kampen lemma in this setting is that there is a sequence of elementary transitions $u = u_0 \to u_1 \to \ldots \to u_n = v$ with respect to the presentation $S = sgp < X : R >$ if and only if there is a (u, v) diagram over this presentation with exactly n regions (bounded faces).

Semigroup diagrams of this type were introduced into the literature by Remmers [96] and have been used by a number of authors to study semigroup presentations. We refer the reader to the book of Higgins [39] for an exposition of some of the uses that have been made of these diagrams - in particular for some results on "small overlap" semigroups and for a version of Remmer's proof of a result of Adian about the embeddability of a semigroup that is left and right "cycle free" in a group.

We turn now to an exposition of some recent (as yet unpublished) work of Kilibarda concerning the algebra of semigroup diagrams. Semigroup diagrams have in general a much simpler structure than group diagrams (van Kampen diagrams). In particular, the fact that every positively labelled path in a semigroup diagram may be extended to a transversal in the sense described above is very useful and imposes severe restrictions on the common boundary of two faces of such a diagram. For example, no edge can be on the left boundary of two distinct faces of a semigroup diagram. Using ideas of this type, one is able to define a process of reduction on semigroup diagrams and a natural multiplication on the set of semigroup diagrams associated with a given presentation.

A pair of faces F_1 and F_2 of a semigroup diagram Δ is called a "reducible pair" if F_1 and F_2 are mirror images of each other; i.e., F_1 has a boundary cycle of the form $\alpha\beta^{-1}$ and F_2 has a boundary cycle of the form $\beta\gamma^{-1}$ where α, β and γ are positively labelled paths and $\phi(\alpha) = \phi(\gamma)$. In this case one can form a new semigroup diagram Δ' by eliminating the common boundary β from Δ and then extracting the new region that is formed from F_1 and F_2 by identifying the paths α and γ. In this case we say that Δ' has been constructed from Δ by one elementary reduction. We define a semigroup diagram Δ to be reduced if it does not have any reducible pair of faces. If Δ and Δ' are two semigroup diagrams over the same presentation we say that they are equivalent if and only if either Δ is identical to Δ' or there is a sequence $\Delta = \Delta_0, \Delta_1, \dots, \Delta_k = \Delta'$ for some k, such that for each $j < k$, one of Δ_{j+1} and Δ_j comes from the other by elementary reduction. It is clear that equivalence of diagrams defined this way is an equivalence relation.

Lemma 2.1 *(Kilibarda) There is exactly one reduced diagram in each equivalence class.*

We denote by $r(\Delta)$ the unique reduced diagram that is equivalent to Δ. The lemma enables us to introduce a natural multiplication on the set of reduced semigroup diagrams associated with a fixed presentation $S = sgp < X : R >$ in the following way. Let Γ be a reduced (u, v) diagram over S and Γ' a reduced (t, w) diagram over S. The product of Γ and Γ' will be defined if and only if $v = t$. In this case we can form a new diagram (a (u, w) diagram over S) by identifying the right boundary of Γ with the left boundary of Γ' and taking the union of the two diagrams under this identification. Denote the new diagram by $\Gamma * \Gamma'$. Then define

$$\Gamma \cdot \Gamma' = r(\Gamma * \Gamma').$$

With respect to this multiplication, the set of reduced semigroup diagrams over S becomes a groupoid. The identities (objects) of this groupoid may be identified with the (u, u) diagrams that have no faces - a reduced (u, v) diagram may be viewed as an arrow (morphism) in this category from the trivial (u, u) diagram to the trivial (v, v) diagram. The maximal subgroups (loop monoids) of this category provide a measure of the essentially different elementary transitions with respect to the presentation that take us from a word u to itself. In particular, the maximal subgroups are all trivial if and only if the presentation is aspherical in the sense of Ivanov [41]. This is a somewhat more restricted notion of asphericity than that of Cho and Pride [21] and may be best thought of as an analogue for semigroups of the notion of "diagrammatic asphericity" in the sense of [19].

The groupoid of semigroup diagrams over the presentation $S = sgp < X : R >$ defined above may be identified with the fundamental groupoid $\pi(K_S)$ of a certain complex K_S. We refer to [40] or [23] or [50] for the notion of the fundamental groupoid of a complex. The complex K_S is defined as follows.

$V(K_S)$: the set of vertices of K_S is X^* (the set of all words over X).

$E(K_S)$: there is an edge e in K_S with initial vertex $\sigma(e) = w_1 \in X^*$ and terminal vertex $\tau(e) = w_2 \in X^*$ whenever there exists a pair of words $x, y \in X^*$ and a relation $r = s$ (or $s = r$) in R such that $w_1 = xry$ and $w_2 = xsy$. The edge e depends not only on r and s but also on the context (x, y) in which they occur. Thus there may be multiple edges between w_1 and w_2 corresponding to application of the same relation but in different contexts. The inverse of the edge e defined above has initial edge w_2 and terminal edge w_1 and is determined by the same relation that was used to define e, in the same context.

$C(K_S)$: there is a 2-cell (face) F in K_S with boundary path $e_1e_2e_3e_4$ whenever there are two disjoint appearances of relators in a word $w \in X^*$, such as $w = xr_1yr_2z$ for some $x, y, z \in X^*$ and some relations $r_1 = s_1$ and $r_2 = s_2$ in R. Then e_1, e_2, e_3 and e_4 are edges with initial and terminal vertices given by:

$$\sigma(e_1) = w = \tau(e_4), \sigma(e_2) = xs_1yr_2z = \tau(e_1),$$

$$\sigma(e_3) = xs_1ys_2z = \tau(e_2), \sigma(e_4) = xr_1ys_2z = \tau(e_3).$$

This 2-cell has an obvious inverse with boundary path $(e_1e_2e_3e_4)^{-1}$.

Theorem 2.2 *(Kilibarda.) Let $S = sgp < X : R >$ be any semigroup presentation. Then the groupoid of semigroup diagrams over S is isomorphic to the fundamental groupoid $\pi(K_S)$ of the complex K_S defined above. In particular, the maximal subgroups of this groupoid of semigroup diagrams are the fundamental groups of K_S.*

REMARKS. (1) Theorem 2.2 may be used to calculate the maximal subgroups of the groupoid of semigroup diagrams over a presentation. For example if $S = sgp < x, y : xy = yx >$ (the free commutative semigroup on two generators) then one sees that all of the corresponding maximal subgroups of the associated groupoid are trivial, so this presentation is aspherical in the sense described above. This observation also essentially follows from the results of Adian [1] or Ivanov [41]. However the free commutative semigroup on more than 2 generators is not aspherical.

(2) It can be shown from Theorem 2.2 that the maximal subgroups of the groupoid of semigroup diagrams over any finitely presented semigroup have decidable word problem. It can also be shown that any finite direct product of finitely generated free groups arises this way. A complete characterization of the class of groups that arise in such a way is at present unknown.

(3) It is interesting to note that a complex isomorphic to K_S has been introduced independently by S.Pride [93] in connection with his work on low dimensional homotopy of monoid presentations.

We turn now to a discussion of the analogue of the Cayley graph of a group presentation in the setting of monoid presentations. For the remainder of this section,

M will denote a monoid presentation of the form $M = mon < X : R >$ where R is a set of monoid relations of the form $u = v$, for some $u, v \in X^*$. Such presentations include the semigroup presentations considered above but we do not now explicitly exclude the case that some of the relations may be of the form $u = 1$ for some $u \in X^+$. In any case we want to view M as a monoid now, even if there are no relations of this form in R. We let π denote the natural map from X^* onto M. One may form the Cayley graph $\Gamma(X : R)$ for such a presentation just as for a group presentation. The vertices of $\Gamma(X : R)$ are the elements of M and there is an edge labelled by x from m to $m\pi(x)$ for each m in M and x in X. We regard $\Gamma(X : R)$ as a word graph over X (i.e., a graph whose edges are labelled over X) in this case, but note that it is not an inverse word graph since there is no "inverse edge" labelled by x^{-1} from $m\pi(x)$ to m.

In order to study the word problem for a monoid presentation M as above, Stephen showed in his thesis [117] that it is convenient to introduce a family of birooted word graphs $B\Gamma(w)$, one for each word w in X^*. The graphs $B\Gamma(w)$ are defined as follows. The underlying graph is the restriction of the Cayley graph $\Gamma(X : R)$ to the set of elements n in M such that $\pi(w)M$ is contained in nM in the monoid M (i.e. n is greater than or equal to $\pi(w)$ in the $R-$class order on M). Thus the set of vertices of $B\Gamma(w)$ is the set of all such elements defined above and the edges are just the edges of $\Gamma(X : R)$ that connect vertices of this form. There are also two distinguished vertices (roots) namely 1 (the initial vertex) and $\pi(w)$ (the terminal vertex). One may view $B\Gamma(w)$ as an automaton over X with states the vertices of the graph and initial state 1 and terminal state $\pi(w)$. Stephen shows in [117] that the language accepted by this automaton is exactly the set of words in X^* that are equivalent to w under the presentation. Thus one solves the word problem for M if one has an effective procedure for constructing the automata $B\Gamma(w)$ (or the language accepted by these automata) for each w in X^*.

There is an iterative construction of the automaton $B\Gamma(w)$ starting from the "linear automaton" of the word w. If w is the word $w = x_1x_2...x_n$ in X^* then the linear automaton of w is the automaton with $n+1$ states $v_0, v_1, ..., v_n$ and an edge labelled by x_i from v_{i-1} to v_i for $i = 1, \ldots, n$. The initial state is v_0 and the terminal state is v_n. Clearly the language accepted by this automaton is just $\{w\}$. Starting with this automaton, one then builds a sequence of automata obtained by applying operations of the following types:

(M1) Expansions: if in an automaton A over X (i.e. a birooted word graph over X) there are two states v_1 and v_2 and a path p from v_1 to v_2 labelled by one side (say u) of a relation $u = v$ in R, but no path from v_1 to v_2 labelled by the other side, then we expand the automaton A by adding a new path from v_1 to v_2 labelled by v. The resulting automaton A' is said to be obtained from A by an expansion.

(M2) Foldings: if in a birooted word graph (automaton) over X there are two edges with the same initial vertex and the same label, then we identify these edges. The resulting automaton A' is said to be obtained from A by folding edges. Note that this is the same as the operation (O3) that was used in the Todd-Coxeter algorithm outlined above, the difference being that in this case we only fold edges with the same initial vertex and the same label, whereas in (O3) we also fold edges with the same terminal vertex and the same label. If we perform all possible foldings to a finite automaton A the result is a deterministic automaton, but not necessarily an injective automaton. By an injective automaton we mean an automaton where every letter induces a partial one-one function.

Stephen has shown [117] that the operations (M1) and (M2) are confluent - i.e., if we start with an automaton A and apply one of these operations to obtain A' and a possibly different operation to A to obtain A'', then there is a third automaton B such that B can be obtained from A' by applying a finite sequence of operations of the form (M1) or (M2) and also B can be obtained from A'' by applying a finite sequence of these operations. The idea then is that we can build $B\Gamma(w)$ by starting with the linear automaton of w and applying operations of the form (M1) or (M2) successively (in any order). The resulting automata become "better and better approximations" to $B\Gamma(w)$ and "in the limit" we obtain $B\Gamma(w)$ this way. These ideas are made precise in [117]. We include some brief explanation of the details below and refer to [117] for more detail. We also refer to [59] for the relevant definitions and concepts from category theory.

In [117] it is shown that the class of birooted word graphs over X is a cocomplete category (see [59] for the definition of this concept) and so in particular this category has colimits. If (α, Γ, β) and $(\alpha', \Gamma', \beta')$ are birooted word graphs obtained from the linear automaton of w by repeated applications of operations (M1) and (M2) and if $(\alpha', \Gamma', \beta')$ is obtained from (α, Γ, β) by one application of one of these operations, then the language accepted by (α, Γ, β) is contained in the language accepted by $(\alpha', \Gamma', \beta')$ and there is a natural morphism of birooted word graphs from (α, Γ, β) to $(\alpha', \Gamma', \beta')$. So there is a natural diagram (directed system) of birooted word graphs obtained from the linear graph of w this way. Stephen shows that $B\Gamma(w)$ is the colimit of this diagram of birooted word graphs in the category of all birooted word graphs over X.

Theorem 2.3 *(Stephen, [117]). Let $M = mon < X : R >$ be a monoid presentation and let w be any word in X^*. Then $B\Gamma(w)$ is the colimit (direct limit) in the category of birooted word graphs over X of the diagram of birooted word graphs that are obtained from the linear graph of w by repeated applications of operations (M1) and (M2) relative to this presentation.*

REMARKS. (1). In order to apply Theorem 2.3 to construct the automata $B\Gamma(w)$ relative to some given presentation, one uses the confluence of the operations (M1) and (M2) to devise a scheme for iteratively applying these operations so that one "eventually" builds in all applications of all possible expansions and all possible foldings. In practice this can be very difficult, especially if the resulting automaton $B\Gamma(w)$ is infinite, and there has not been extensive use of this technique in the literature to date.

(2). In his paper [63], McCammond constructs automata somewhat similar to the automata $B\Gamma(w)$ to solve the word problem for the free semigroups in the Burnside variety of semigroups defined by an identity of the form $x^a = x^{a+b}$ for $a > 5$. However the automata that McCammond uses are essentially non - deterministic, unlike the automata $B\Gamma(w)$ of Stephen.

2.3. Inverse Monoids

Inverse monoids arise naturally whenever one is concerned with injective partial functions (i.e., partial one-one functions on a set). The class of inverse monoids forms a variety of algebras of type $< 2, 1, 0 >$ defined by associativity and the identities

$$x = xx^{-1}x, \ (x^{-1})^{-1} = x, \ x \cdot 1 = 1 \cdot x = x, \ xx^{-1}yy^{-1} = yy^{-1}xx^{-1}.$$

This last law expresses the fact that idempotents commute and is one of the crucial properties of inverse monoids. As a consequence of this, the set $E(M)$ of idempotents of an inverse monoid forms a semilattice with respect to the multiplication, which may be extended to the natural partial order on M by defining $a \leq b$ iff $a = eb$ for some e in $E(M)$. Inverse monoids have been considered extensively in the literature. We refer to the book of Petrich [88] for much basic information about inverse monoids and in particular for various ways in which they may be constructed from groups and semilattices.

The study of algorithmic problems in inverse semigroups may be traced back to the work of Gluskin [33] who studied the structure of the free inverse monoid on one generator, but it was not until the work of Munn [75] and Scheiblich [110] that the structure of the free inverse monoid on a set X was known and the word problem for the free inverse monoid was solved. Subsequently, Stephen [117], [118] developed a framework for studying the word problem for presentations of inverse monoids. Stephen's work built on the work of Munn [75] and Margolis and Meakin [54], where graphical methods akin to those commonly used in combinatorial group theory were introduced. This work has led to a substantial amount of subsequent work by Margolis, Meakin, Stephen and others on the word problem for inverse monoids [57], [55], [11], [114], [43] and varieties of inverse monoids [58], [68]. Many other results on varieties of inverse monoids, properties of the free inverse monoid and the elementary theory of the free inverse monoid have appeared in the literature. Closed inverse submonoids of the free inverse monoid have been studied by Margolis and Meakin [56] and applications of these ideas to the study of subgroups of free groups have been developed in [56] and [12]. These ideas will be discussed in the next section of this paper. Much of this work on algorithmic problems in inverse semigroups has been discussed in the survey article of Meakin [66], so we shall include only a brief discussion of part of this theory here.

The structure of the free inverse monoid $FIM(X)$ on a set X can be elegantly described via the Cayley graph (tree) of the free group $FG(X)$ on X. We present this as a special case of a more general construction of Margolis and Meakin [54] here, since we will need the ideas of this more general construction in Section 6.7 of this paper. Let $G = gp < X : R >$ be a group presentation and let $\Gamma(X : R)$ be the corresponding Cayley graph. We build an inverse monoid from $\Gamma(X : R)$ in the following way. Define $M(X : R) = \{(\Gamma, g) : \Gamma$ is a finite connected subgraph of $\Gamma(X : R)$ containing the vertices 1 and g of $\Gamma(X : R)\}$ with multiplication

$$(\Gamma, g) \cdot (\Delta, h) = (\Gamma \cup g \cdot \Delta, gh),$$

where $g \cdot \Delta$ denotes the translate of Δ on the left by g in $\Gamma(X : R)$ and $\Gamma \cup g \cdot \Delta$ simply denotes the union of the corresponding subgraphs of $\Gamma(X : R)$. In [54] it is shown that $M(X:R)$ is an inverse monoid with maximal group homomorphic image G and that the passage from G to $M(X:R)$ defines a functor from the category of $X-$generated groups to the category of $X-$generated "E-unitary" inverse monoids that is a left adjoint to the maximal group image functor. We recall that an inverse monoid M is said to be E-unitary if the natural map σ from M onto its maximal group image is "idempotent-pure" (i.e., the inverse image of 1 under σ is just $E(M)$).

The monoids $M(X:R)$ constructed above have many pleasant properties. In particular, if $G = gp < X : \phi > = FG(X)$, then the corresponding inverse monoid $M(X:\phi)$ is isomorphic to the free inverse monoid $FIM(X)$ on X. Thus the elements of $FIM(X)$ may be thought of as birooted inverse word trees of the form $(1,\Gamma,g)$ where Γ is a finite subtree of the tree $\Gamma(X:\phi)$ containing 1 and g, and the multiplication is as specified above. If w is a word in $(X\cup X^{-1})^*$ then the birooted word tree associated with w is $(1, MT(w), r(w))$, where $r(w)$ is the reduced form of w in the usual group theoretic sense and $MT(w)$ is the "Munn tree" of w - i.e., $MT(w)$ is the subtree of $\Gamma(X:\phi)$ traversed when one reads the word w in $\Gamma(X:\phi)$, starting at the vertex 1 and ending at the vertex $r(w)$. Munn's solution to the word problem in $FIM(X)$ is: if u and v are two words in $(X\cup X^{-1})^*$, then $u=v$ in $FIM(X)$ if and only if $MT(u)=MT(v)$ and $r(u)=r(v)$.

The Munn tree $MT(w)$ of a word w may be regarded as a finite inverse automaton over $X\cup X^{-1}$ in the usual way (with initial state 1 and terminal state $r(w)$). The vertices of this graph (automaton) are the elements of $FIM(X)$ that are related to w via Green's R-relation on $FIM(X)$ and the language accepted by this automaton is the set of words v in $(X\cup X^{-1})^*$ such that $v\geq w$ in the natural partial order on $FIM(X)$. These ideas have been extended by Stephen [117], [118] to study arbitrary presentations of inverse monoids.

We denote by $inv < X : T >$ the inverse monoid presented by the set X of generators and the set T of relations. Here we may consider T as a set of relations of the form $u_i = v_i$ where u_i and v_i are elements of the free monoid $(X\cup X^{-1})^*$ and we interpret $inv < X : T >$ as the image of $(X\cup X^{-1})^*$ obtained by imposing the relations in T together with all of the identities that define the variety of inverse monoids: alternatively, one may view the u_i and v_i as elements of $FIM(X)$ and $inv < X : T >$ as the image of $FIM < X >$ obtained by imposing these relations. It follows immediately by universal considerations that $gp < X : T >$ is the maximal group homomorphic image of $inv < X : T >$.

There is a well developed and useful analogue of the Cayley graph for studying presentations of inverse monoids, but no analogue at the present time of the methods of van Kampen diagrams or semigroup diagrams discussed above. We briefly indicate the situation with respect to the Cayley graph. One may define the Cayley graph $\Gamma(X:T)$ of an inverse monoid presentation in the obvious way, just as for semigroup presentations or group presentations. It turns out to be useful in this setting, however, to restrict this Cayley graph to the set of vertices that are $R-$related to the word under consideration.

Let $M = inv < X : T >$ and let w be any word in $(X\cup X^{-1})^*$. Denote the natural map from $(X\cup X^{-1})^*$ onto M by π. We define a word graph $S\Gamma(w)$ over $X\cup X^{-1}$ as follows. The vertices of $S\Gamma(w)$ consist of the elements m of M that are related via Green's $R-$relation to $\pi(w)$. There is an edge labelled by x from

m to $m\pi(x)$ for each x in $X \cup X^{-1}$ provided both m and $m\pi(x)$ are $R-$related to $\pi(w)$ in M. This implies that there is also an edge labelled by x^{-1} from $m\pi(x)$ to m in $S\Gamma(w)$, which we take as the inverse of the original edge. Thus $S\Gamma(w)$ is in fact an inverse word graph over X in the sense discussed above. The birooted inverse word graph $A(w) = (\pi(ww^{-1}), S\Gamma(w), \pi(w))$ may be regarded as an inverse automaton in the usual way (with initial state $\pi(ww^{-1})$ and terminal state $\pi(w)$). This automaton has been referred to as the "Schutzenberger automaton of w" in the literature and the underlying graph $S\Gamma(w)$ as the "Schutzenberger graph of w" since $S\Gamma(w)$ is the graph of the Schutzenberger representation of M relative to the $R-$class in consideration. Note the distinction between $S\Gamma(w)$ and the graph $B\Gamma(w)$ constructed above in connection with a monoid presentation. Note also that if $M = FIM(X) = inv < X : \phi >$ then for each word w in $(X \cup X^{-1})^*$, there is an isomorphism between $A(w)$ and the birooted inverse word graph $(1, MT(w), r(w))$.

In his work [118], Stephen shows that the language accepted by $A(w)$ is in fact the set of words v in $(X \cup X^{-1})^*$ such that $v \geq w$ in the natural order on M and that two words u and v in $(X \cup X^{-1})^*$ are equal in M if and only if $A(u)$ and $A(v)$ accept the same language, i.e., if and only if $A(u)$ and $A(v)$ are isomorphic as birooted word graphs over X. Thus we have a solution to the word problem for M if we have an effective procedure for constructing each automaton $A(w)$ for each word w in $(X \cup X^{-1})^*$. Stephen goes on to provide an iterative procedure for constructing the automata $A(w)$. We briefly describe this procedure below.

We again start with the linear automaton of the word w, as described in the previous section. However we now regard this automaton as an inverse automaton over X (i.e., we associate with each edge labelled by x in $X \cup X^{-1}$ an inverse edge labelled by x^{-1} in the usual way). Let us denote the resulting linear automaton of w by $LinA(w)$. Note that w is in the language accepted by $LinA(w)$, but that this language is infinite if $w \neq 1$ (for example, $(ww^{-1})^n w$ is in this language for each $n \geq 1$). It is not difficult to see that the language accepted by $LinA(w)$ is contained in the language accepted by $A(w)$, as described above. Starting with $LinA(w)$, one builds a sequence of automata obtained by applying operations of the following types:

(I1) Expansions: this is exactly the same as the expansion operation (M1) discussed in the previous section. However we build in the additional requirement that each edge with label x from $X \cup X^{-1}$ in the new automaton must come equipped with its natural inverse edge labelled by x^{-1}.The resulting automaton satisfies condition (P1) of section 2.1.

(I2) Foldings: this is exactly the same as the folding operation (O3) discussed in section 2.1. Note that if we apply all possible foldings to a finite automaton that satisfies condition (P1) we obtain an inverse automaton. Note the distinction between this operation and the folding operation (M2) of section 2.2.

Once again the effect of applying an expansion or a folding to an automaton that satisfies (P1) is to obtain another such automaton which accepts all words in the language of the original automaton and there is a morphism from the first birooted word graph to the second. Thus one may again consider the diagram (directed system) of all birooted word graphs obtained from the linear automaton $LinA(w)$ by successive applications of the expansion and folding operations (I1) and (I2), and A(w) is the colimit (direct limit) of this diagram of automata.

Theorem 2.4 *(Stephen, [118]) Let $M = inv < X : T >$ be an inverse monoid presentation and let w be any word in $(X \cup X^{-1})^*$. Then $A(w)$ is the colimit (direct limit) in the category of birooted word graphs over $X \cup X^{-1}$ of the diagram (directed system) of birooted word graphs that are obtained from the linear graph $LinA(w)$ of w by repeated applications of operations (I1) and (I2) relative to this presentation.*

Note that the operations (I1) and (I2) are again confluent in this setting, so that one may proceed to construct $A(w)$ from $LinA(w)$ by applying a sequence of expansions and foldings in any desired order, so long as one eventually builds in all possible expansions and foldings. In practice, one usually arranges that all intermediate automata in such a sequence are inverse automata, by applying all possible foldings to the automaton under consideration. Note also that the Todd-Coxeter algorithm for constructing the Cayley graph of a finitely presented group is in fact a special case of this general construction of Stephen, corresponding to the presentation $M = inv < X : R, xx^{-1} = x^{-1}x = 1$ for all $x \in X >$.

The iterative construction of the automata $A(w)$ associated with an inverse monoid presentation $M = inv < X : T >$ has proved to be a powerful tool in analysing such monoids. In his original paper and thesis [117], [118], Stephen examines several examples where these methods enable him to solve the word problem for classes of inverse monoid presentations that could not be handled effectively using the usual "linear" arguments on words that are often used in semigroup theory. We shall not cite a list of all other instances in the literature where these methods have been used, but we refer the reader to the following papers where very significant use of these methods has been made:

(a) the paper [55] where the word problem is solved for inverse monoids presented by finitely many relations of the form $e_i = f_i$, where e_i and f_i are idempotents of the free inverse monoid;

(b) the paper [58] where the word problem for the free semigroups in the Burnside variety of inverse semigroups defined by the identity $[x^a = x^{a+b}]$ for $b \leq a$ is solved. This problem is surprisingly easy to solve, quite in contrast to the situation for Burnside varieties of semigroups or groups (see [45] for references and discussions);

(c) the paper [94] where free combinatorial strict inverse semigroups are studied.

(d) the paper [43] where these methods are used to study the structure of free products of inverse semigroups;

(e) the paper [68] where undecidability of the word problem for the Mal'cev product of the variety of semilattices and the variety of abelian groups is proved: this result will be discussed in more detail in Section 6;

(f) the paper [24] where wreath products of varieties of inverse semigroups are studied;

(g) the paper [11] where it is shown that the word problem for one relator inverse monoids of the form $M = inv < X : e = 1 >$ (for e an idempotent of $FIM(X)$) can be solved in polynomial time: this is in contrast to the general situation for the inverse monoids considered in (a) above, where the complexity of the solution to the word problem appears very high.

It is perhaps worth mentioning here that the word problem for one relator inverse monoids of the form $M = inv < X : w = 1 >$ remains unsolved. Even in the case where w is a reduced word, it can be proved that a positive solution to this problem would imply a positive solution to the one relator semigroup problem for semigroups of the form $S = sgp < X : u = v >$. The situation here is in sharp contrast to the situation for one relator groups, where the word problem was solved by Magnus in the early 1930's and one relator monoids of the form $M = mon < X : w = 1 >$ where the word problem was solved by Adian [1]. Some recent (unpublished) work of Ivanov, Margolis and Meakin shows that there may be a reasonable chance of settling the word problem for inverse monoids of the form $M = inv < X : w = 1 >$ in the case where w is *cyclically* reduced. In this case it is possible to use the van Kampen diagrams of combinatorial group theory to show that M is E-unitary. We refer the reader to the paper [57] for some early ideas about this problem.

3. Immersions, Inverse Automata and Algorithmic Problems for Subgroups of Free Groups

3.1. Preliminaries

The notion of an *immersion*, that is a locally injective graph morphism, has been used to prove a number of results about free groups [32], [116]. We show here that inverse monoids play the same role in the theory of immersions that groups play in the theory of coverings. Just as a cover over the bouquet of X circles is "essentially" a representation of the free group $FG(X)$ by permutations, we will see that an immersion over the bouquet of X circles is "essentially" a representation of the free inverse monoid $FIM(X)$ by partial one to one maps. We make this precise by showing that the category of immersions over the bouquet of X circles is naturally equivalent to the category of representations of $FIM(X)$.

By picking two distinguished vertices to be regarded as a start and terminal state we can think of immersions as inverse automata. If the terminal state coincides with the start state then the subset of $FIM(X)$ accepted by this inverse automaton is a closed inverse submonoid. This means that this submonoid is a filter in the natural partial order on $FIM(X)$ $(x \leq y \leftrightarrow x = xx^{-1}y)$. Conversely, every closed inverse submonoid is accepted by such an inverse automaton. Under this identification, finitely generated closed inverse submonoids correspond to finite inverse automata. Furthermore, we will see that finitely generated closed inverse submonoids are precisely the closed inverse submonoids of finite index in the sense of inverse semigroup theory [88] and that they are also exactly the closed submonoids that are rational subsets of $FIM(X)$ in the sense of formal language theory.

We can also associate an inverse automaton with every subgroup H of the free group $FG(X)$. This construction, to be explained below, is usually more "compact" than the corresponding permutation automaton (that is, the cover corresponding to H). In this way, just as subgroups of finite index correspond to finite covers, that is finite permutation automata, finitely generated subgroups correspond exactly to finite inverse automata. This observation then gives a quick proof of many finiteness results for subgroups of free groups. Furthermore, this allows us to use algebraic properties of the transition (inverse) monoid of an inverse automaton to give a "syntactic" classification of finitely generated subgroups of free groups. We exploit this in two ways. First of all, we develop an analogue of Eilenberg's Theorem that gives a bijection between pseudo-varieties of finite monoids and semigroups and classes of regular languages called varieties of languages [27], [49], [90].

Secondly, we show that certain algorithmic questions about subgroups of free groups can be shown to be polynomial equivalent to corresponding properties of inverse automata and their transition monoids. This allows us to prove for example, that the problem of testing whether a finitely generated subgroup is pure or p-pure for a given prime p is PSPACE-complete.

Finally we return to inverse semigroup theory by showing how to construct all closed inverse submonoids of a free inverse monoid by looking at actions of groups on trees. For detailed proofs of the material presented here, see [56], [12].

3.2. Covers are Permutation Automata

In this section we recall some basic definitions and establish the connection between covers of graphs and permutation automata. All our graphs will be in the sense of [113]. Let Γ' and Γ be graphs. A morphism $\eta : \Gamma' \to \Gamma$ is a *cover* if for all vertices $v' \in \mathrm{Vert}(\Gamma')$, η induces a bijection:

$$\eta_{v'} : Star_{v'} \to Star_{v'\eta}$$

where $Star_v = \{e \in \ \mathrm{Edge}(\Gamma) \mid e \text{ starts at } v\}$. A permutation automaton over an alphabet X is an automaton over $X \cup X^{-1}$ such that each $x \in X \cup X^{-1}$ induces a permutation on the state set and such that the inverse letter induces the inverse permutation. For this and the next section our automata do not have any initial state or terminal states specified. There is an evident notion of morphism of cover, morphism of automata and morphism of representation of FG(X).

Theorem 3.1 *The following categories are naturally equivalent:*

a) The category of (connected) covers over the bouquet of X circles.

b) The category of (transitive) representations of FG(X) by permutations.

c) The category of (connected) permutation automata over X.

It is well known that a transitive representation of $FG(X)$ is equivalent (in the category of representations of $FG(X)$) to the coset representation of $FG(X)$ modulo the stabilizer of any vertex. In this way, finite connected covers correspond exactly to conjugacy classes of subgroups of $FG(X)$ of finite index. However, finitely generated subgroups need not correspond to finite covers. We shall see later that they do correspond to finite immersions and can be classified by finite inverse monoids.

3.3. Immersions are Inverse Automata

In this section we show that immersions and inverse automata play the same role as covers and permutation automata in the previous section. Let Γ' and Γ be graphs. A morphism $\eta : \Gamma' \to \Gamma$ is a *immersion* if for all vertices $v' \in \mathrm{Vert}(\Gamma')$, η induces an injection:

$$\eta_{v'} : Star_{v'} \to Star_{v'\eta}.$$

There is an evident notion of morphism of immersion, morphism of inverse automata and morphism of representation of FIM(X).

Theorem 3.2 *[56] The following categories are naturally equivalent:*
(a) The category of (connected) immersions over the bouquet of X circles.
(b) The category of (transitive) representations of FIM(X) by injective functions.
(c) The category of (connected) inverse automata over X.

Just as transitive representations of $FG(X)$ are determined up to equivalence by subgroups of $FG(X)$ a theorem of Schein [111] shows that transitive representations of $FIM(X)$ are determined up to equivalence by closed inverse submonoids of $FIM(X)$. That is, the stabilizer of any vertex in a transitive representation of $FIM(X)$ is a closed inverse submonoid of $FIM(X)$. Conversely, any closed inverse submonoid N of $FIM(X)$ determines a transitive representation of $FIM(X)$ on the so called ω-cosets of N. Two closed inverse submonoids of $FIM(X)$ determine equivalent representations (in the category of representations) if and only if they are conjugate in an appropriate sense. See [88] for more details. Thus up to conjugacy, closed inverse submonoids and immersions are equivalent notions. We shall see that unlike the case of groups, finite immersions correspond to both finite index and finitely generated closed inverse submonoids.

3.4. Some Classical Applications of Immersions

In this section we review a number of classical applications of immersions. Many of these have appeared in various places in the literature and in many guises. We claim no originality here, but just gather these together for purposes of illustration.

3.4.1. *The Free Group is Residually Finite*

Let $w \neq 1$ be an element of $FG(X)$. We wish to construct a finite group H and a morphism $\phi : FG(X) \to H$ such that $w\phi \neq 1$. There are many well known constructions of such an H. Here is a simple construction based on inverse automata.

Step 1 Let $w \neq 1 \in FG(X)$. Construct the linear inverse automaton $LinA(w)$ that reads the word w (see Section 2.3). $LinA(w)$ has a state set Q_w with length(w) + 1 states.

Step 2 Since $LinA(w)$ has a finite number of states, we can complete the automaton to a permutation automaton P_w on the same set Q_w. Since the word w does not fix the initial state, the transition induced by w in the transition group H of P_w is not equal to 1 in H.

EXAMPLE

Let $X = \{a, b\}$ and $w = aba^{-1}b^{-1}a$. Then the automaton $LinA(w)$ is shown in Figure 1 below where 1 is the initial state, 6 is the terminal state.

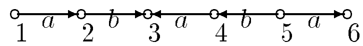

Fig. 1. $LinA(w)$

Now we can define P_w to be **any** permutation automaton that extends the action of $LinA(w)$ on the state set $\{1, \ldots, 6\}$. Then since the image of w in the transition monoid H of P_w maps 1 to 6, w does not induce the identity transformation and thus the image of w in H is not the identity.

3.4.2. *The Generalized Word Problem for $FG(X)$*
The generalized word problem for a finitely presented group G is to decide given an element $g \in G$ and a finite subset $Y \subseteq G$ whether or not $g \in < Y >$ where $< Y >$ is the subgroup generated by Y. It is well known that this problem is solvable for the free group $FG(X)$ on a finite set X. One method is to use linear methods related to Neilsen reduction to try to represent g as a product of free generators for $< Y >$. Here we show how to associate a finite inverse automaton $\mathcal{A}(< Y >)$ with $< Y >$ that accepts the unique reduced word representing g if and only if $g \in < Y >$. In [56] it is shown that this automaton just depends on the group $< Y >$ and not the particular generating set Y assuming each element of Y is given by its reduced representative.

ALGORITHM

Input: A finite set $Y \subset FG(X)$ and an element $w \in FG(X)$.

Output: "Yes" if and only if $w \in < Y >$.

Method

Step 1 Construct the flower automaton $\mathcal{F}(Y)$ of Y. $\mathcal{F}(Y)$ has a state i that is the unique initial and terminal state and one "petal" for each $y \in Y$ that spells out y in a loop from i to i.

Step 2 $\mathcal{F}(Y)$ may not be an inverse automaton, so we fold edges (in the sense of Section 2.1) until an inverse automaton is obtained. It can be shown that this process leads to a unique inverse automaton $\mathcal{A}(< Y >)$ that depends only on the subgroup generated by Y and not the set Y itself.

Step 3 Answer "Yes" if and only if the reduced word representing w labels a path from the initial state to itself. It can be shown that a reduced word labels such a path if and only if it is in $< Y >$, so that this algorithm is correct.

It is not difficult to see that $\mathcal{A}(< Y >)$ can be constructed in polynomial time from the input. Furthermore, it is easy to see that testing whether the reduced word representing w is accepted in $\mathcal{A}(< Y >)$ is also done in polynomial time. Thus this algorithm is polynomial. Avenhaus and Madlener [5] studied the complexity of the generalized word problem and proved that this problem is P-complete with respect to logspace reductions.

3.4.3. *Marshall Hall's Theorem*
In [36], Hall proved that every finitely generated subgroup of a free group is closed in the profinite topology of the free group. Since this conference proceedings contains a number of articles on the profinite topology and its relationship to automata theory and semigroup theory, we will not repeat definitions here. We will just call a subgroup H of a group G *closed* if H is the intersection of some collection of subgroups of G each of which has finite index.

We can use the automaton $\mathcal{A}(< Y >)$ constructed in the last section to prove Hall's Theorem. Let Y be a finite set contained in $FG(X)$ and let $w \in FG(X)$ be an element that is not in $< Y >$. Consider $\mathcal{A}(< Y >)$. Since w is not in $< Y >$, either w does not label a loop in $\mathcal{A}(< Y >)$ at the start state or w does not label any path in $\mathcal{A}(< Y >)$. In the first case, let $\mathcal{A}_w(< Y >) = \mathcal{A}(< Y >)$. In the second case, let $\mathcal{A}_w(< Y >)$ be the inverse automaton obtained from $\mathcal{A}(< Y >)$ by sewing on a path reading w from the initial state (adding new states when the path "falls off" $\mathcal{A}(< Y >)$).

In either case, $\mathcal{A}_w(< Y >)$ is a finite inverse automaton with the property that w labels a path starting at the initial state and ending elsewhere. Once again, we complete $\mathcal{A}_w(< Y >)$ to a permutation automaton on its state set in any possible way. We obtain a finite permutation automaton $\mathcal{P}_w(< Y >)$ with the same property. Now let H_w be the stabilizer of the initial state in $\mathcal{P}_w(< Y >)$. Then H_w has the following properties:

- H_w has finite index in $FG(X)$.
- w is not in H_w.
- $< Y > \subseteq H_w$.

It follows that $< Y >$ is the intersection of the collection $\{H_w \mid w$ is not in $< Y >\}$. Since each of these has finite index we have proved that $< Y >$ is closed.

3.4.4. *The Finite Index Problem*
The finite index problem asks whether a finitely generated subgroup of a group has finite index. Again the inverse automaton can be used to solve this problem. The equivalence of the first two conditions of the next theorem appears in [116], but has been known for some time. Let $\mathcal{I}(< Y >)$ be the transition monoid of $\mathcal{A}(< Y >)$. $\mathcal{I}(< Y >)$ is a finite inverse monoid that once again is independent of Y.

Theorem 3.3 *Let Y be a finite subset of $FG(X)$ and $H =< Y >$. Then the following conditions are equivalent:*

a) H has finite index.

b) $\mathcal{A}(H)$ is a permutation automaton.

c) $\mathcal{I}(H)$ is a finite group.

It is easy to see that one can test if a finite inverse automaton $\mathcal{A}$ is a permutation automaton in polynomial time in terms of the number of edges of $\mathcal{A}$. It follows that the finite index problem can be solved in polynomial time.

Theorem 3.3 forms a basis for our ideas expounded below. We wish to translate properties of finitely generated subgroups into properties of the corresponding **finite** inverse automaton and **finite** inverse monoid. If the corresponding property can be effectively checked in the finite automaton or finite inverse monoid, then we have essentially described an algorithmic solution to the original problem. If the problem just depends on the inverse automaton, then since we can construct the automaton in polynomial time from a generating set for the subgroup, we have efficiently reduced the problem in group theory to the corresponding problem in automaton theory. This is the case in Theorem 3.3 above. If on the other hand, the corresponding problem depends on the structure of the inverse monoid, then the complexity of the problem may increase significantly. This is because the size of the inverse monoid may be exponential in the size of the input. We will see this phenomenon in the next section.

3.5. Purity of Subgroups: A PSPACE-complete problem

In this section we use the ideas outlined above to show that the problem of detecting whether a finitely generated subgroup of $FG(X)$ is pure or p-pure is PSPACE-complete. This is the first "natural" problem on free groups that we know of that is PSPACE-complete.

Recall that a subgroup H of a group G is *pure* (*p-pure* for a given prime p) if for all $g \in G$, $g^n \in H$ for some $n > 0$ (n relatively prime to p), implies that $g \in H$.

Theorem 3.4 *[12] Let H be a finitely generated subgroup of $FG(X)$. Then the following conditions are equivalent:*

a) H is pure (p-pure, for given prime p).

b) Every subgroup of $\mathcal{I}(H)$ is trivial (a p-group).

It is straightforward to see that given a finite inverse automaton $\mathcal{A}$, then checking whether every subgroup in the transition monoid of $\mathcal{A}$ is trivial or a p-group can be done in polynomial space. Thus Theorem 3.4 gives a PSPACE algorithm for detecting whether a finitely generated subgroup of $FG(X)$ specified by a finite set of generators is pure or p-pure. We note that the conditions in Theorem 3.4 depend on the structure of the transition monoid of $\mathcal{A}(H)$. Thus this condition is not necessarily detectable in polynomial time from the input. Note that in the finite index problem of the preceding section, the condition that the transition monoid of $\mathcal{A}(H)$ be a group is equivalent to the easily checkable property that $\mathcal{A}(H)$ itself be a permutation automaton. We can ask whether the property of testing whether every subgroup of the transition monoid of a finite inverse automaton $\mathcal{A}$ is trivial or a p-group is reducible to some easily checkable property of the automaton itself.

The main theorem of [12] shows that these problems are in fact PSPACE-complete so it is unlikely that any of these problems has a polynomial time solution. This is done by showing that these problems are polynomial equivalent to certain problems about inverse automata. A number of interesting problems about inverse automata are needed and are also shown to be PSPACE-complete.

In particular, the following problems are considered in [12] :

The intersection emptiness problem Given a collection of finite injective or inverse automata, $\mathcal{A}_i$, $1 \leq i \leq n$, is there a word accepted by all of the $\mathcal{A}_i$, $1 \leq i \leq n$?

The aperiodicity problem Given an injective or inverse automaton $\mathcal{A}$, is every subgroup of the transition monoid of $\mathcal{A}$ trivial?

The generation problem Given an inverse automaton with state set Q and a test injective function $f : Q \to Q$, is f a member of the transition monoid of the automaton?

The p-periodicity problem Given an injective or inverse automaton $\mathcal{A}$ and a prime number p, is every subgroup of the transition monoid of $\mathcal{A}$ a p-group?

Theorem 3.5 *[12] The problems:*

a) the intersection emptiness problem,

b) the aperiodicity problem,

c) the generation problem,

d) the p-periodicity problem,

e) the purity of subgroup problem,

f) the p-purity of subgroup problem.

are all polynomial time reducible to each other. All of these problems are PSPACE-complete.

To do this we make use of injective Turing machines and a theorem of Bennett [9]. These problems were known to be PSPACE-complete for finite automata in general [48]. [20]. On the other hand, for group automata, these problems have very fast parallel algorithms and are in the class NC [6]. Thus, with respect to complexity theory, inverse semigroups behave more like arbitrary semigroups than groups.

3.6. Pseudo-varieties of Finite Inverse Monoids and Closed Inverse Submonoids of Free Inverse Monoids

We have seen that we have the following equivalences for a finitely generated subgroup H of a free group $FG(X)$.

- H is finitely generated if and only if $\mathcal{I}(H)$ is a finite inverse monoid.
- H has finite index if and only if $\mathcal{I}(H)$ is a finite group.
- H is pure if and only if $\mathcal{I}(H)$ is a finite inverse monoid with trivial subgroups.
- H is p-pure if and only if $\mathcal{I}(H)$ is a finite inverse monoids and all subgroups in $\mathcal{I}(H)$ are p-groups.

These correspondences suggest that the properties of subgroups of free groups that can be detected by these methods are related to varietal properties of the corresponding inverse monoid. This is in total analogy with the relationship between pseudo-varieties of finite semigroups and monoids and varieties of languages in the sense of Eilenberg. See [27], [49], [90] for an extensive introduction to these ideas.

Ruyle [99] has developed a similar correspondence between pseudo-varieties of finite inverse monoids and pseudo-varieties of finitely generated closed inverse submonoids of free inverse monoids. He also has developed a number of correspondences between classes of finitely generated subgroups of free groups and pseudo-varieties of finite inverse monoids, but the connection is somewhat more subtle, due to the fact that there are many immersions that represent a subgroup of the free group. That is, the inverse automaton $\mathcal{A}(H)$ that we have associated with a subgroup H of $FG(X)$ has the minimal number of states among all immersions representing H as a stabilizer of a vertex. We can obtain all other such immersions by sewing on trees at points on $\mathcal{A}(H)$ that have no edge reading some letter in $X \cup X^{-1}$. The resulting transition monoid of such an inverse automaton can have varietal properties that can be surprisingly different form the transition monoid of $\mathcal{A}(H)$ itself. Depending on the application, it is sometimes useful to consider not only $\mathcal{A}(H)$ but one or all of these larger immersions corresponding to H.

In the cases listed above the corresponding properties are independent of which inverse automaton we consider. For example, one inverse automaton representing H has transition monoid with trivial subgroups if and only if all inverse automata representing H have this property. Details and more examples can be found in [99].

3.7. Structure of Closed Inverse Submonoids of Free Inverse Monoids

The ideas and applications of the preceding sections were geared towards applications to the classification of subgroups of free groups. In this section we return internally to the free inverse monoid itself and show how immersions can be used to give a structure theorem for closed inverse submonoids of $FIM(X)$. We will see that every such object is constructed from the free action of a (free) group on a tree.

Let T be a tree and let G be a group acting freely on the left of T. This means that the stabilizer of each vertex of T is trivial. It is well known that G is a free group and that every free group has such an action on a tree[113].

Fix a vertex $v \in V(T)$. Let $M(T,G,v) = \{(t,g) | t$ is a finite subtree of $T, g \in G$ and $v, gv \in V(t)\}$ and define a multiplication by

$$(t_1, g_1)(t_2, g_2) = (t_1 \cup g_1 t_2, g_1 g_2).$$

Here $g_1 t_2$ denotes the translate of the tree t_2 and $t_1 \cup g_1 t_2$ is the subtree whose set of vertices (edges) is the union of those of the trees t_1 and $g_1 t_2$.

Example Let $G = FG(X)$ and let $T = \Gamma(X)$ be the Cayley graph of G relative to the usual presentation. Then T is a tree and G acts freely on T by left multiplication. By Munn's theorem, [75], $M(T,G,1)$ is isomorphic to $FIM(X)$, the free inverse monoid on X. Clearly in this case we obtain the construction of $FIM(X)$ given in Section 2.3.

One of the main theorems of [56] states that monoids of the form $M(T,G,v)$ are precisely the class of closed inverse monoids of free inverse monoids. We review the main constructions here.

Let N be a closed inverse submonoid of $FIM(X)$. Let Γ_N be the immersion corresponding to N. That is Γ_N is the automaton of right ω-cosets of N and N is represented as the stabilizer of (the right ω-coset) N.

Let T_N be the universal covering graph of Γ_N. It is well known that T_N is a tree and that the fundamental group $G = \pi_1(\Gamma_N)$ acts freely on the left of T_N (by deck transformations). Furthermore, T_N embeds into the Cayley graph of the free group. We can arrange that 1 is a vertex in this embedded image and that 1 covers N in the covering map from T_N to Γ_N. Finally it is known that Γ_N is the quotient of T_N under the action of G which can be represented as a subgroup of $FG(X)$.

Lemma 3.1 *[56] Let N be a closed inverse submonoid of $FIM(X)$. Then the maximal group image of N is isomorphic to $G = \pi_1(\Gamma_N)$. Furthermore, N is isomorphic to $M(T_N, G, 1)$.*

Thus, every closed inverse submonoid N of $FIM(X)$ can be naturally constructed form the topological invariants T_N and $G = \pi_1(N)$. Conversely, every monoid constructed this way is isomorphic to a closed inverse submonoid of an appropriate free inverse monoid.

Lemma 3.2 *[56] Let G be a group acting freely on a tree T with root v. Let $\Gamma = G\backslash T$ and let X be an orientation of Γ. Then $M(T,G,v)$ is isomorphic to a closed inverse submonoid of $FIM(X)$.*

It is not true that every closed inverse submonoid of $FIM(X)$ is a free inverse monoid. For example, the semilattice of idempotents of $FIM(X)$ is closed, but not a free inverse monoid. In [56] it is proved that every closed inverse submonoid has a retraction onto a free inverse monoid. This is proved by translating the topological notion of contracting a spanning tree into the algebraic setting of inverse semigroups. This and many other facts concerning these monoids can be found in [56]. It is also shown in [56] how to use the free inverse category on a graph Γ to classify immersions over Γ.

3.8. Finiteness Conditions

In this section, we show that the closed inverse submonoids of free inverse monoids satisfy finiteness properties not shared by subgroups of free groups. This allows one to lift properties of finitely generated subgroups of free groups to the closed inverse submonoid generated by the same set. Results such as Howson's Theorem follow easily.

Recall that the set of rational subsets of a monoid M is the smallest collection of subsets of M containing the singletons and closed under finite union, product of subsets and submonoid generation (i.e. Kleene star). A subset S of M is recognizable if there is a finite monoid N and a morphism $f : M \to N$ and a subset P of N such that $S = Pf^{-1}$. See [10] for details. Let $Rat(M)$ be the set of rational subsets of M and let $Rec(M)$ be the set of recognizable subsets of M. We have the following important theorems.

Theorem 3.6 *(Kleene) If M is a finitely generated free monoid, then $Rec(M) = Rat(M)$.*

Theorem 3.7 *(Anissimov and Seifert). Let G be a finitely generated group and let H be a subgroup of G. Then $H \in Rec(G)$ if and only if H has finite index in G. $H \in Rat(G)$ if and only if H is finitely generated.*

It follows from Theorem 3.7 that if G is any infinite group, then the trivial subgroup is rational, but not recognizable. We also list the following consequence of Kleene's theorem due to McKnight.

Theorem 3.8 *Let M be a finitely generated monoid. Then Rec(M) is contained in Rat(M).*

We say that a closed inverse submonoid N of $FIM(X)$ is *finitely generated* if N is the smallest closed inverse submonoid containing a finite set Y. That is $N = \{n \in FIM(X) | n \geq y_1^{\epsilon_1}, \ldots, y_m^{\epsilon_m} y_i \in Y, \epsilon_i \in \{1, -1\}\}$. N has *finite index* if N has a finite number of ω-cosets. That is, the immersion Γ_N corresponding to N is finite or equivalently by Theorem 3.2, N is recognized by a finite inverse automaton. Putting this altogether we have the following theorem proved in [56].

Theorem 3.9 *[56] Let M = FIM(X) and let N be a closed inverse submonoid of M. Then the following conditions are equivalent:*

a) N is recognized by a finite inverse automaton.

b) N has finite index in M.

c) N corresponds to a finite immersion over the bouquet of circles.

d) N is a recognizable subset of M.

e) N is a rational subset of M.

f) N is finitely generated.

As an application, we can obtain a quick proof of Howson's Theorem: the intersection of two finitely generated subgroups of a free group is also finitely generated. Indeed, if $N_i \subset FG(X)$ is generated by a finite set Y_i, $i = 1, 2$, let $\hat{N}_i$ be the closed inverse submonoid of $FIM(X)$ generated by Y_i. Then $\hat{N}_i$ is recognized by a finite inverse automaton $\mathcal{A}_i$ by Theorem 3.9. A standard construction of the theory of automata allows us to construct a finite automaton recognizing $\hat{N}_1 \cap \hat{N}_2$. Thus, $\hat{N}_1 \cap \hat{N}_2$ is a finitely generated closed inverse submonoid of $FIM(X)$ and so is its image $N_1 \cap N_2$ in $FG(X)$.

4. Burnside-type Properties, Symbolic Dynamics, Identities and Quasi-Identities of Finite Semigroups

We now turn to algorithmic problems in varieties.

First of all we need some basic definitions.

- A semigroup S is *periodic* if all its one-generated subsemigroups are finite, equivalently if for every element $x \in S$ there exist two different numbers m_x and n_x such that $x^{m_x} = x^{n_x}$.
- A nil-semigroup is a semigroup in which a power of every element is equal to zero.
- A nilpotent semigroup of degree n is a semigroup where any product of n elements is zero. Every finite nil-semigroup is nilpotent [22].
- A variety is called periodic if it consists of periodic semigroups. A variety is called non-periodic if it contains a non-periodic semigroup, or, equivalently, if it contains the additive semigroup of natural numbers.

4.1. Burnside Problems in Varieties of Semigroups

As we mentioned in the introduction, the description of varieties where periodic semigroups are locally finite (see Theorems 4.4 and 4.5 below) plays an exceptional role in the study of algorithmic problems in semigroup varieties. Most of the results about algorithmic problems in varieties would be impossible to obtain without it.

The first result about Burnside-type problems in semigroup varieties was published by Morse and Hedlund [74]. The result was the following:

Theorem 4.1 *There exist an infinite semigroup with three generators that satisfies the identity*[1] $x^2 = 0$ *and an infinite semigroup with two generators that satisfies the identity* $x^3 = 0$.

Morse and Hedlund used certain infinite words W_1 and W_2 over a 2-letter alphabet and a 3-letter alphabet, which *avoid* the words x^3 and x^2 respectively. In general if u is a word and ϕ is an endomorphism of a free semigroup then $\phi(u)$ is called a *value* of u. A word u is called *avoidable* by a word W if W does not contain any values of u. A word u is called *avoidable* if it is avoided by an infinite word over a finite alphabet.

We present the Thue construction of the word W_1 in Section 6.5 below.

There is a natural correspondence between infinite words over a finite alphabet and finitely generated semigroups (see Section 4.6). This correspondence implies the following connection between the avoidability of words and Burnside-type properties (see [8]).

Theorem 4.2 *A word u is avoidable if and only if the variety given by the identity $u = 0$ is not locally finite.*

The next step was made by D.B.Bean, A.Ehrenfeucht, and G.McNulty [8], and independently by Zimin [125]. They found algorithms for checking if a word is avoidable.

[1] We use the short expression $u = 0$ for the pair of identities $ux = u, xu = u$ where x does not occur in u.

We will need the definition of the words Z_n, which we will call Zimin words:

$$Z_1 = x_1, \dots, Z_{n+1} = Z_n x_{n+1} Z_n.$$

The following theorem is a translation of results from [8] and [125] into the language of varieties.

Theorem 4.3 *Let $\mathcal{V}$ be a variety of semigroups given by a(possibly infinite) set of identities $\{u = 0 \mid u \in \Sigma\}$. Assume that the number of variables occurring in words of Σ is n. Then the following conditions are equivalent.*

1. *$\mathcal{V}$ is locally finite.*
2. *Z_n contains a value of some word u in Σ.*

This theorem gives an algorithmic description of locally finite varieties defined by identities of the form $u = 0$. Notice that all these varieties consist of nil-semigroups.

The next result by Sapir [102] gives an algorithmic description of arbitrary varieties where nil-semigroups are locally finite.

Theorem 4.4 *[102] Let $\mathcal{V}$ be a variety of semigroups given by a (possibly infinite) set of identities Σ. Assume that the number of variables occurring in words of Σ is n. Then the following conditions are equivalent.*

1. *All nil-semigroups from $\mathcal{V}$ are locally finite.*
2. *All semigroups from $\mathcal{V}$ satsfying the identity $x^2 = 0$ are locally finite.*
3. *There exists an identity $u = v \in \Sigma$ such that Z_{n+1} contains a value $\phi(u)$ for some endomorphism ϕ but $\phi(u) \neq \phi(v)$.*

The next theorem, also from [102], describes varieties where all periodic semigroups are locally finite in the class of varieties with "good" groups and in the class of non-periodic varieties.

Theorem 4.5 *[102] Let $\mathcal{V}$ be a variety of semigroups given by a (possibly infinite) set of identities Σ. Assume that the number of variables occurring in words of Σ is n. Assume also that the variety $\mathcal{V}$ either is non-periodic or contains no non-locally finite groups of finite exponent. Then the following conditions are equivalent.*

1. *All periodic semigroups from $\mathcal{V}$ are locally finite.*
2. *All nil-semigroups from $\mathcal{V}$ are locally finite.*
3. *There exists an identity $u = v \in \Sigma$ such that Z_{n+1} contains a value $\phi(u)$ for some endomorphism ϕ but $\phi(u) \neq \phi(v)$.*

These two theorems have many interesting corollaries and applications (see [102], [104]). Let us present just one of them immediately. Others will be discussed later.

Theorem 4.6 *[102] A finitely based periodic semigroup variety is locally finite if and only if all its groups and all its nil-semigroups are locally finite.*

4.2. Burnside Properties and Identities of Semigroups

One of the central problems in unversal algebra is the following problem connecting the syntactic and semantic ways of describing varieties.

Tarski's Finite Basis Problem Is the set of finite algebras that generate finitely based varieties recursive?

For groups the answer is given by the celebrated theorem of Oates and Powell.

Theorem 4.7 *(Oates-Powell) Every finite group G generates a finitely based variety.*

The same result is not true for semigroups. Let B_2^1 denote the semigroup consisting of the 2×2 matrix units together with the 0 matrix and the identity matrix.

Theorem 4.8 *(Perkins [86]) The variety of semigroups generated by B_2^1 is not finitely based.*

Theorem 4.4 allows one to get a much stronger result. Indeed, one can reformulate Theorem 4.4 in the following way:

Theorem 4.9 *(Sapir, [102]) Let $\mathcal{V}$ be a variety where every nil-semigroup is locally finite. Suppose $\mathcal{V}$ does not satisfy any non-trivial identity of the form $Z_n = W$. Then $\mathcal{V}$ cannot be defined by a finite number of identities.*

This theorem turned out to be a very powerful tool in studying the finite basis property in semigroup varieties.

For example one can easily establish that the Brandt monoid B_2^1 does not satisfy any non-trivial identity of the form $Z_n = W$. This immediately implies the following theorem.

Theorem 4.10 *([102]) A locally finite variety of semigroups is not finitely based provided it contains the Brandt monoid B_2^1.*

A finite algebra which cannot belong to a locally finite finitely based variety is called *inherently non-finitely based* [87]. Theorem 4.10 actually states that B_2^1 is an inherently non-finitely based semigroup. Theorem 4.10 answered a question by G.McNulty and C.Shallon [65].

It is clear that if an inherently non-finitely based algebra A divides a finite algebra B (that is A is a homomorphic image of a subalgebra of B) then B itself is inherently non-finitely based . Thus if B_2^1 divides a finite semigroup S, then S is not finitely based. Using a result of Volkov [115], it turns out that the converse is also true for inverse semigroups.

Theorem 4.11 *([102]) Let S be a finite inverse semigroup. Then considered as a semigroup the following conditions are equivalent.*

a) S is finitely based.

b) S is weakly finitely based.

c) B_2^1 does not divide S.

Since it is decidable whether B_2^1 divides a finite inverse semigroup S, given the multiplication table of S, it is decidable whether a given inverse semigroup is finitely based as a semigroup. The Tarski problem for arbitrary finite semigroups remains open.

In [103] Sapir gave the following algorithmic description of all finite inherently non-finitely based semigroups.

Theorem 4.12 *([103]) A finite semigroup S of order n is inherently non-finitely based if and only if it does not satisfy any non-trivial identity of the form $Z_{n^2} = W$.*

We say that a finite algebra A is *weakly finitely based* if A belongs to some locally finite finitely based variety (that is if A is not inherently non-finitely based). Of course, every finitely based finite algebra is weakly finitely based, but we will see that the opposite implication does not hold.

4.3. Identities of Inverse Semigroups

B_2^1 is also an inverse semigroup. It is easy to see that if $s \in B_2^1$, then s^{-1} is the transpose matrix of s. As we mentioned in Section 2, the class of inverse semigroups is a variety of algebras of type $< 2, 1 >$ consisting of multiplication and inversion.

We can thus study the variety of inverse semigroups generated by B_2^1. Notice that B_2^1 plays a very important role in the theory of varieties of inverse semigroups. In particular by a result of E.Kleiman [47] every inverse semigroup variety that does not contain B_2^1 and that is generated by a finite semigroup, is finitely based. He also proved the following result.

Theorem 4.13 *(E.Kleiman [47]) The variety of inverse semigroups generated by B_2^1 is not finitely based.*

Given Theorem 4.11 and Kleiman's Theorem 4.13, one is lead to expect that B_2^1 is inherently non-finitely based even when considered as an inverse semigroup. Thus the following result is very surprising.

Theorem 4.14 *([109]) Let S be a finite inverse semigroup. Then, considered as an inverse semigroup, S is weakly finitely based.*

That is, every finite inverse semigroup belongs to a locally finite finitely based variety of inverse semigroups. It is a major open problem to decide which inverse semigroups are finitely based when considered as inverse semigroups. It is still possible that every inverse semigroup having B_2^1 as a divisor is not finitely based as an inverse semigroup.

Theorem 4.14 is a consequence of the following two theorems. Let Z_n' be the prefix of Z_n consisting of all but the last letter. We will say that a semigroup S has height $\leq h$ if and only if every descending chain of principal right ideals has length $\leq h$.

Theorem 4.15 *Every inverse semigroup of height $n-2$ satisfies the identity $Z_n' = Z_n' x_1 x_1^{-1}$.*

Theorem 4.16 *Let S be a finitely generated inverse semigroup satisfying the identity $Z_n' = Z_n' x_1 x_1^{-1}$. Let all subgroups of S be locally finite. Then S is finite.*

For example, by Theorem 4.15, B_2^1 belongs to the variety defined by the identities $x^2 = x^3$ and $xyxzxyxtxyxzxy = xyxzxyxtxyxzxyxx^{-1}$. From Theorem 4.16, it follows that the variety defined by these two identities is locally finite.

4.4. Quasi-varieties of Semigroups

Quasi-varieties of algebras are classes defined by quasi-identities, that is formulas of the following form:

$$(\forall x_1, \ldots, x_n)\ u_1 = v_1 \& \ldots \& u_m = v_m \rightarrow u = v \qquad (1)$$

where u_i, v_i, u, v are terms in variables $x_1, \ldots, x_n$.

After varieties, quasi-varieties are the most widely studied classes of algebras. Again, they are studied from both a syntactic and a semantic point of view.

Syntactic A quasi-variety is a class of algebras satisfying a given set of implications.

Semantic A quasi-variety is a class of algebras closed under taking subalgebras, products and "ultraproducts". (See [18] for a definition of ultraproducts).

For example let $Q = \{S | (xz = yz) \rightarrow (x = y)\}$. Q is the quasi-variety of right cancellative semigroups.

There are the evident notions of finitely generated, finitely based, weakly finitely based and inherently non-finitely based quasi-varieties. Ol'shanskii [81] described the collection of finitely generated quasi-varieties of groups that are finitely based.

Theorem 4.17 *(Ol'shanskii [81]) Let G be a finite group. Then the quasi-variety generated by G has a finite basis of implications if and only if every Sylow subgroup of G is Abelian.*

Sapir [100] has studied the question of which finite semigroups have a finite basis of quasi-identities. In particular, he described finite semigroups without two-sided ideals that have this property. But the general problem remains open. The following theorem is the appropriate analogue to Theorem 4.14.

Theorem 4.18 *(Margolis and Sapir [60]) Let S be a finite semigroup. Then S belongs to a locally finite quasi-variety, defined by a finite set of implications. That is, every finite semigroup is weakly finitely based with respect to quasi-identities.*

Theorem 4.18 is a consequence of the following two theorems.

Theorem 4.19 *Every finite semigroup of height $n-2$ satisfies the quasi-identity $(tZ_n = wZ_n) \to (tZ'_n = wZ'_n)$.*

Theorem 4.20 *Let S be a finitely generated periodic semigroup satisfying the quasi-identity $(tZ_n = wZ_n) \to (tZ'_n = wZ'_n)$. Let all subgroups of S be locally finite. Then S is finite.*

For example, by Theorem 4.19, B_2^1 (as a semigroup) belongs to the quasi-variety defined by the identity $x^2 = x^3$ and the implication

$$\begin{aligned} txyxzxyxuxyxzxyx &= wxyxzxyxuxyxzxyx \to \\ & \quad txyxzxyxuxyxzxy = wxyxzxyxuxyxzxy. \end{aligned}$$

From Theorem 4.20, it follows that this quasi-variety is locally finite.

4.5. Semigroups and Inverse Semigroups

It is hard not to notice a similarity between Theorems 4.15 and 4.19. The reason for this similarity is that these theorems express the same structural property of semigroups in two different cases: the case of inverse semigroups and the case of semigroups. In the first case this property may be expressed by an identity in the second case it can be expressed by a quasi-identity.

Let $u, v \in X^+$. We say that S satisfies the condition $u\mathcal{R}^*v$ if for every homomorphism $\phi : X^+ \to S$, $\phi(u)$ and $\phi(v)$ generate the same principal right ideal in some semigroup T containing S. It is known that in the case of inverse semigroups T may be always taken equal to S. The relation $\mathcal{R}^*$ has been studied extensively by J.Fountain [30].

The following lemmas show us how to express the condition "S satisfies $u\mathcal{R}^*v$" in terms of identities for inverse semigroups and quasi-identities for semigroups. Lemma 4.2 appeared in [84].

Lemma 4.1 *Let S be an inverse semigroup. Then S satisfies the condition $u\mathcal{R}^*v$ if and only if S satisfies the identity $uu^{-1} = vv^{-1}$.*

Lemma 4.2 *Let S be a semigroup. S satisfies the condition $u\mathcal{R}^*v$ if and only if S satisfies the bi-implicaton $tu = wu \leftrightarrow tv = wv$, where t and w are variables not appearing in u and v.*

From these lemmas, it is easy to derive the following corollaries.

Corollary 4.1 *Let S be an inverse semigroup. Then S satisfies $Z_n\mathcal{R}^*Z'_n$ if and only if S satisfies the identity $Z'_nx_1x_1^{-1} = Z'_n$.*

Corollary 4.2 *Let S be a semigroup. Then S satisfies $Z_n\mathcal{R}^*Z'_n$ if and only if S satisfies the quasi-identity $(tZ_n = wZ_n) \to (tZ'_n = wZ'_n)$.*

Now the following theorem implies both Theorems 4.15 and 4.19.

Theorem 4.21 *Let S be a semigroup of height $n-2$. Then it satisfies*

$$Z_n\mathcal{R}^*Z'_n$$

4.6. Semigroups and Symbolic Dynamics

The proofs of Theorems 4.4, 4.16 and 4.20 have one important detail in common. All these proofs employ the following connection between semigroups and symbolic dynamical systems. This connection appeared first in the paper of Sapir [102].

A *symbolic dynamical system* is a closed subset of the Tikhonov product $X^{\mathbf{Z}}$, where X is a finite set with the discrete topology, which is stable under the shift homeomorphism T (this homeomorphism shifts every sequence from $X^{\mathbf{Z}}$ one position to the right).

The correspondence between semigroups and symbolic dynamics is the following. Let $S = < X >$ be an infinite finitely generated semigroup (the same argument may be applied for any universal algebra). Then there is an infinite set T of words over X such that every element of S represented by a word of T cannot be represented by words over X of shorter length. Such words will be called *geodesic words*. These words label geodesics in the Cayley graph of the semigroup. It is clear that every subword of a geodesic word is also a geodesic word. Now, in every word of T, mark a letter which is closest to the center of this word. There must be an infinite subset T_1 of T of words which have the same marked letters, an infinite subset T_2 of T_1 of words which have the same subwords of length 2 containing the marked letters, $\ldots$, an infinite subset T_n of T_{n-1} of words which have the same subwords of length n containing the marked letters, and so on. Therefore there is an infinite word W such that every subword of W is a subword of a word from T. Thus every subword of W is geodesic. Infinite words with this property will also be called *geodesic*. The set $D(S)$ of all infinite geodesic words is a symbolic dynamical system because it is stable under the shift (obviously) and is closed in the Tikhonov topology (this can be easily proved). Conversely, with every symbolic dynamical system D one can associate a semigroup $S(D)$ as follows: $S(D)$ consists of all finite subwords of infinite words from D, and 0. If u and v belong to $S(D)$ then $u \cdot v$ is equal to uv if uv belongs to $S(D)$, or 0 otherwise. It is easy to show that $S(D)$ is a semigroup. The following theorem shows that every symbolic dynamical system is equal to $D(S)$ for some semigroup S.

Theorem 4.22 *(Sapir, [45]) For every symbolic dynamical system D we have*

$$D(S(D)) = D.$$

The correspondence between semigroups and symbolic dynamical systems allows one to show that some important properties of the theory of semigroups and important properties of the theory of symbolic dynamical systems are in fact equivalent. For example, *if a finitely presented semigroup S is periodic then the symbolic dynamical system $D(S)$ does not have cyclic trajectories. The semigroup S is infinite if and only if $D(S)$ is not empty, etc.*

One of the important concepts of the theory of symbolic dynamical systems is the concept of a uniformly recurrent word. An infinite word U is called *uniformly recurrent* if for every finite subword u of U there exists a number $N_U(u)$ such that every subword of U of length $N_U(u)$ contains u as a subword. It is an easy corollary from [31] (see [102] for details) that for every infinite word U there exists a uniformly recurrent word U' such that every subword of U' is a subword of U. It is easy to see that if U belongs to $D(S)$ then U' also belongs to $D(S)$.

Therefore for every infinite finitely generated semigroup $S =< X >$ there exists a uniformly recurrent geodesic word over X (see [102]). Uniformly recurrent words are much more convenient than arbitrary infinite words.

4.7. Uniformly Recurrent Words and Burnside Properties

The following theorem was first proved in [102].

Theorem 4.23 *Let $\alpha \in A^Z$ be uniformly recurrent . Suppose that $a \in A$ and that $\alpha = \alpha_1 a \alpha_2$ is an occurrence of a in α where α_1 is infinite to the left and α_2 is infinite to the right. Then there is a substitution $\phi : \{x_1, x_2, \dots\} \to A^+$ and a sequence of natural numbers $A(n, \alpha)$ such that:*

a) $\phi(x_1) = a$

b) $\alpha_1 a = \alpha_3 \phi(Z_n)$ for some α_3 infinite to the left.

c) $|\phi(Z_n)| \leq A(n, \alpha)$.

This theorem shows that uniformly recurrent words have strong regularity properties. In order to prove that a finitely generated semigroup S is finite, we must be able to translate these regularities into properties about the structure of S.

For example, let us sketch the proof of Theorems 4.16 and 4.19. Let us take a periodic semigroup S with all subgroups locally finite which have the property

$$Z_n' \mathcal{R}^* Z_n. \tag{2}$$

Suppose that S is infinite. Then there exists a geodesic uniformly recurrent word α. Let $A = A(n, \alpha)$ be the number from Theorem 4.23.

Lemma 4.3 *Let u, v, and w be consecutive subwords of α such that u and v have length $\geq A$, w may be empty. Then $Z_n' \mathcal{R}^* Z_n$ implies $uvwu\mathcal{R}^* uvwuv$.*

Proof. Suppose that p is the longest prefix of v such that the relation

$$uvwu\mathcal{R}^* uvwup$$

follows from (2). If $v = p$ then we are done so suppose $v \neq p$. Let $v = paq$ for some letter a and word q. We have $uvwup = upaqwup$. Since upa is a subword of α and $|upa| \geq A$ by Theorem 4.23 there exists an endomorphism ϕ of the free semigroup such that $upa = u_1\phi(Z_n)$ and $\phi(x_1) = a$. Therefore $up = u_1\phi(Z_n')$. Thus $Z_n' \mathcal{R}^* Z_n$ implies $up\mathcal{R}^* upa$. Therefore we have $uvwu\mathcal{R}^* uvwup\mathcal{R}^* uvwupa$ which contradicts the choice of the prefix p. The lemma is proved.

Let us consider the (finite) set of all subwords of α of length $2A$. Since α is uniformly recurrent there exists a number B such that every subword of length B of α contains every subword of this set. Let u and v be any two subwords of α of length $\geq 2B$, let a and b be the corresponding elements in S. An induction similar to that used in Lemma 4.3 gives us the following lemma.

Lemma 4.4 *a and b are $\mathcal{J}$-related, i.e $a = c_1bd_1, b = c_2ad_2$ for some elements c_1, c_2, d_1, d_2 from S.*

Notice that so far we did not use the fact that all subgroups of S are locally finite and that U is geodesic. Thus we actually proved the following fact:

Lemma 4.5 *Let $S =< X >$ be a semigroup satisfying $Z_n'\mathcal{R}^*Z_n$. Let U be any uniformly recurrent word over X. Then there exist only finitely many $\mathcal{J}-$classes of S which contain elements represented by subwords of U.*

Now if we take a uniformly recurrent geodesic α, we can find a natural number N such that all subwords of α of length $\geq N$ will represent elements from the same $\mathcal{J}$-class of S. Readers familiar with the basic structure theory of semigroups [22] will see that the condition $Z_n'\mathcal{R}^*Z_n$ makes S semisimple. They will then know that if all subgroups of S are locally finite then these subwords will represent only a finite number of distinct elements of S. It follows that α is not irreducible and thus S is finite.

Thus we see that if a semigroup satisfies the condition (2) then it satisfies the descending chain condition for ideals generated by factors of a uniformly recurrent sequence. The ordinary descending condition for ideals may not hold in such a semigroup (see [109] for a counterexample).

Applications of uniformly recurrent words are very effective, but are not constructive. Indeed, there is no algorithm to find the number $N_U(u)$. The proofs of Theorems 4.4 and 4.5 have been made constructive in [107] where the analogue of the restricted Burnside problem for semigroup varieties is discussed.

5. Varieties with Decidable Word Problem

5.1. Commutative Semigroups

We say that the word problem is decidable (solvable) in a variety $\mathcal{V}$ of semigroups if it is decidable (solvable) in every semigroup which is finitely presented in $\mathcal{V}$.

The first non-trivial (that is non-locally finite) variety of semigroups with solvable word problem was found independently by A.I.Mal'cev [52] and by Ceitin and Emelichev [28]. It was the variety of all commutative semigroups.

Theorem 5.1 *The variety of all commutative semigroups has a solvable word problem.*

Mal'cev showed that every finitely generated commutative semigroup is faithfully representable by matrices over a suitable field, and every finitely generated ring of matrices is residually finite. Therefore every finitely generated commutative semigroup is residually finite and the McKinsey algorithm (see Theorem 5.6 below) gives the solution to the word problem.

As was pointed out by Emelichev in [28] the proof may be easily deduced from an old paper by Hermann [38] devoted to the membership problem for ideals in the ring of polynomials. Indeed, let $S =< X|u_1 = v_1, \ldots, u_r = v_r >$ be any finitely presented commutative semigroup. Emelichev [28] proved that a relation $u = v$ holds in S if and only if the polynomial $p = u - v$ belongs to the ideal of the ring of polynomials $\mathbf{Q}[X]$ generated by polynomials $p_i = u_i - v_i$. This, in turn, is equivalent to the solvability of the following equation over $\mathbf{Q}[X]$:

$$p_1 f_1 + p_2 f_2 + \ldots + p_r f_r = p \tag{3}$$

with unknowns $f_1, \ldots, f_r$. Now we can use the following result from [38]:

Theorem 5.2 *Let $d = max\{deg(p_1), deg(p_2), \ldots, deg(p_r)\}$. If the equation (3) has a solution then there is a solution with $deg(f_i) \leq deg(p) + (rd)^{2^{|X|}}$.*

It is clear that if there is a bound for the degrees of the unknown polynomials in (3) then the number of coefficients of these polynomials is also bounded. Then (3) is equivalent to a finite system of linear equations over the field of rational numbers, which can be solved by, say, the Gauss elimination algorithm.

Ballantyne and Lankford used this connection between commutative semigroups and ideals in the ring of polynomials to apply the Gröbner basis method to the word problem in commutative semigroups [7].

A proof of Theorem 5.1 based on other ideas is presented below in Section 5.3.1.

Taiclin [119] proved the following result which is much stronger than Theorem 5.1.

Theorem 5.3 *The elementary theory of every finitely presented commutative semigroup is decidable.*

5.2. The Description of Varieties of Semigroups with Solvable Word Problem

The first example of a finitely based proper variety of semigroups with an unsolvable word problem is the variety of Murskii [76]. This was the only known example until 1983 when a deep study of varieties with decidable word problem was initiated by I.Mel'nichuk. In particular, she proved [70] that the word problem is undecidable in any variety which contains a non-locally finite variety of semigroups given by identities of the form $u = 0$. Such varieties were described by Bean, Ehrenfeucht, McNulty and Zimin (see Theorem 4.3 above). She also proved that the word problem is decidable in any finitely based variety which satisfies the permutation identity $x_1 x_2 \ldots x_n = x_{\sigma(1)} x_{\sigma(2)} \ldots x_{\sigma(n)}$ where σ is a permutation of the symbols $1, 2, \ldots n$.

Then Mel'nichuk, Sapir, and Kharlampovich [71] found a minimal variety with an undecidable word problem, a boundary between decidability and undecidability. This was the variety generated by the semigroup S_2 from Section 6.4 below. The semigroup S_1 and its dual semigroup $\overleftarrow{S_1}$ from Section 6.4 also appeared in [71].

Finally, Sapir [101], [105] proved that S_1 and $\overleftarrow{S_1}$ also generate boundaries between decidability and undecidability, proved that there are no more boundaries among non-periodic varieties of semigroups, that every periodic semigroup in a non-periodic variety with decidable word problem must be locally finite, and that every non-periodic variety with an undecidable word problem contains one of these three varieties. He also showed that a periodic variety with a solvable word problem and locally finite groups must be locally finite itself. This is everything that one can hope to get, because the problem of describing non-locally finite periodic varieties of groups with solvable word problem is hopeless. It also turned out that many other conditions for finitely presented semigroups in varieties are equivalent to the solvability of the word problem in this variety.

Some of the results from [101] and [105] of Sapir are summarized in the following two theorems.

We will need the following two finite subsemigroups of B_2:

$$P = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\},$$

$$T = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$

For every semigroup S the dual (anti-isomorphic) semigroup is denoted by $\overleftarrow{S}$, the semigroup S with an identity element adjoined is denoted by S^1. For every variety $\mathcal{V}$ let $FP(\mathcal{V})$ be the class of all semigroups which are finitely presented in $\mathcal{V}$.

Theorem 5.4 *[105] Let $\mathcal{V}$ be a finitely based non-periodic variety. Then the following conditions are equivalent.*

1. *The word problem is decidable in any semigroup from $FP(\mathcal{V})$.*
2. *The elementary theory is decidable for any semigroup from $FP(\mathcal{V})$.*
3. *Every semigroup from $FP(\mathcal{V})$ is residually finite.*
4. *Every semigroup from $FP(\mathcal{V})$ is representable by matrices over a field.*
5. *Every semigroup from $FP(\mathcal{V})$ is Hopfian.*

A. *Every nil-semigroup from $\mathcal{V}$ is locally finite and $\mathcal{V}$ does not contain any of the semigroups $\overleftarrow{P} \times P^1$ or $\overleftarrow{P}^1 \times P$ or T.*

B. *Every nil-semigroup of $\mathcal{V}$ is locally finite and for some natural numbers k, m, n, p the variety $\mathcal{V}$ satisfies one of the following identities:*

$$x^n y (z^k t^k)^p z^m = x^m (t^k x^k)^p y z^n, \quad n = m + kp; \tag{4}$$

$$x y^n z = y^k x y^m z y^p, \quad n > m. \tag{5}$$

Theorem 5.5 *[105] Let $\mathcal{V}$ be a finitely based periodic variety of semigroups in which all groups are locally finite. Then each of the conditions 1, 2, 3, 4, from Theorem 5.4 is equivalent to the condition that $\mathcal{V}$ is locally finite.*

Let us analyze these theorems. From Theorems 5.4, 5.5 it follows that in the class of non-periodic semigroup varieties where all nil-semigroups are locally finite every variety which has an undecidable word problem must contain one of the three finite semigroups:

$$\overleftarrow{P} \times P^1, \ \overleftarrow{P}^1 \times P, \ T. \tag{6}$$

On the other hand, every non-periodic variety, containing one of these three semigroups, has an undecidable word problem. Therefore the following three varieties are the only minimal varieties with undecidable word problem among non-periodic varieties with locally finite nil-semigroups:

$$\mathrm{var}(\overleftarrow{P} \times P^1 \times \mathbf{N}),\ \mathrm{var}(\overleftarrow{P}^1 \times P \times \mathbf{N}),\ \mathrm{var}(T \times \mathbf{N})$$

where $\mathbf{N}$ is the semigroup of natural numbers with respect to addition. In fact, these three varieties coincide with the varieties generated by the semigroups S_1, $\overleftarrow{S}_1$, and S_2 respectively, presented in Section 6.4 below.

It is easy to see that each of these varieties is a join of a locally finite variety (generated by a finite semigroup) and the variety of all commutative semigroups. Both have solvable word problem. Therefore we have *examples of varieties with unsolvable word problem that are joins of varieties with solvable (and even polynomially solvable) word problem.*

If a finitely based variety does not contain any of the three semigroups (6) but still has an unsolvable word problem then either it contains a non-locally finite nil-semigroup or it is periodic and contains a "bad" group, a non-locally finite group of finite exponent (see Theorem 5.4). Theorems 4.4 and 4.5 show that there are no minimal finitely based non-locally finite varieties of nil-semigroups, that is every non-locally finite finitely based variety of nil-semigroups contains a proper subvariety with the same properties. Thus in the first case the variety does not contain any minimal variety with unsolvable word problem. In the second case the situation is more complex. We do not know any minimal periodic variety of groups with unsolvable word problem. Moreover an analogy between nil-semigroups and periodic groups make the following two conjectures of Sapir seem reasonable:

Conjecture 5.1 *Every finitely based periodic group variety with undecidable word problem contains a proper subvariety with an undecidable word problem.*

Conjecture 5.2 *Every non-locally finite finitely based variety of periodic groups has an undecidable word problem.*

Of course, we do not have any hope to prove these conjectures. So far we have only a few examples of periodic group varieties with undecidable word problem (see Section 6.6).

Thus, modulo groups, there are exactly three minimal varieties of semigroups with an undecidable word problem.

5.3. Methods of Proving Decidability of The Word Problem in a Variety

The first and most universal method of proving the decidability of the word problem uses the connection between residual finiteness and the solvability of the word problem discovered by Mal'cev [52].

Theorem 5.6 *Let $\mathcal{V}$ be a finitely based variety of algebras. Let A be an algebra finitely presented in $\mathcal{V}$. If A is residually finite then the word problem in A is decidable.*

In order to solve the word problem in a residually finite algebra Mal'cev used an algorithm which was essentially due to McKinsey [64].

As we have mentioned above, in many cases (see, for example, the previous section) a variety has decidable word problem only if all algebras finitely presented in this variety are residually finite. But this is not always so. For example the group variety $\mathcal{N}_2\mathcal{A}$, which is the Mal'cev product of the variety of nilpotent groups of class two and the variety of Abelian groups, has a decidable word problem [44] but the following theorem holds.

Theorem 5.7 *[46] There exists a non-residually finite group finitely presented in the variety $\mathcal{N}_2\mathcal{A}$.*

Thus one needs other methods of proving the decidability of the word problem. There are, of course, methods that apply when we are solving the word problem in an algebra given by a specific system of defining relations. For example the word problem is decidable in the case when one can find a terminating Church-Rosser presentation of a finitely presented semigroup or group S. Then there exists a "canonical" form for every word over the alphabet of generators and an effective procedure which transforms every word to its canonical form.

It is very unusual, however, that a terminating Church-Rosser presentation is known for an arbitrary algebra finitely presented in a variety. In connection with the role that is played by the Church-Rosser method in computer algebra [17], it would be very interesting to describe varieties where every finitely presented algebra has a Church-Rosser presentation.

An example is the variety of commutative semigroups where terminating Church-Rosser presentations were found by Ballantyne and Lankford [7].

Experience shows that in general, in order to solve the word problem in an arbitrary finitely presented algebra (semigroup, group, etc.) A in a variety, one has to use methods which are in a sense opposite to the "canonical form" methods described above. Instead of finding a "canonical" word in the set of words which are equal to a given word u one has to consider this set as a whole and find some "hidden" structure on this set. To illustrate this idea we will consider two examples: a variety of semigroups and a variety of groups.

5.3.1. *Semigroups*

Here we would like to present some ideas for proving the decidability of the word problem in varieties of semigroups, employed initially in [13], and later in [71], [69], and [105]. To illustrate these ideas we will show how they work in the variety of commutative semigroups. Other methods of proving the decidability of the word problem in this variety have been discussed in Section 5.1.

First of all recall that the free n-generated commutative semigroup A_n with an identity element is simply a direct product of n semigroups of natural numbers. Therefore every element of this semigroup may be represented by a vector of natural numbers[2]. We can consider A_n as a partially ordered set with the natural coordinate-wise order. It is clear that $u < v$ in A_n if and only if u divides v. It is easy to prove

[2] Zero also is a natural number.

that every subset M of A_n has minimal elements, and every element of M is greater than or equal to a minimal element. The first part of the following simple statement is attributed to Dickson [26]. The second part is a constructive version of Dickson's result and is also well known.

Lemma 5.1 *Every infinite set T of elements in A_n contains two comparable elements. If no two elements in a set $T \subset A_n$ are comparable and T contains an element with all coordinates $\leq s$ then $|T| \leq n!s^n$.*

Let R be a finite set of defining relations, i.e. a finite subset of $A_n \times A_n$. We want to show that the word problem is decidable in the factor semigroup of A_n modulo the congruence generated by R. In other words, we want an algorithm, which, given a pair of elements (u, v) in $A_n \times A_n$, says if u equals v modulo the relations of R.

Take an element u in A_n. We will describe the set $M(u)$ of all elements of A_n which are equal to u modulo R.

Let M_{min} be the set of minimal elements in $M(u)$. Every element v in $M(u)$ is greater than or equal to an element w from M_{min}. Hence $v = we$ for some e. Since both v and w are equal to u modulo R we have that $u = ue$ (Mod R). Such an element e is called a *unit* for u. Let $E(u)$ be the set of all units. Then $E(u)$ is a subsemigroup of A_n and is closed under taking quotients, that is if $e_1 f = e_2$ and e_1, e_2 are units then f is also a unit. This implies easily that the semigroup $E(u)$ is generated by the subset E_{min} of its minimal elements. Notice also that if e is a unit for u and $w \in M(u)$ then, of course, $we \in M(u)$.

By Lemma 5.1 both sets M_{min} and E_{min} are finite. Therefore we have the following description of $M(u)$:

$$M(u) = \{(\prod_{e \in E_{\mathrm{min}}} e^{k_e})w \mid w \in M_{\mathrm{min}}, k_e \in \mathbf{N}\}.$$

This description would give us a solution of the word problem, if we had a process of finding the sets M_{min} and E_{min}. This process is almost straightforward. We simply apply the relations from R to u until we find all elements from M_{min} and E_{min}. Of course, the most tricky thing in such processes is a stop sign: the sign which shows that we have found all elements that we need, and we can stop and relax. This is organized as follows.

Denote the maximal length of words from R by $\ell(R)$. If M is a subset of A_n then let $T(M)$ be the set of all elements of A_n which can be obtained from elements of M by applying relations of R (at most one application for each relation and for each element of M).

Let us construct a sequence of sets M_n. Let $M_0 = \{u\}$. Suppose we have constructed the set M_n. Let E_n be the set of all quotients of elements from M_n.

Let us apply T to M_n many times until we obtain all elements of the form

$$(\prod_{e \in \mathrm{min}(E_n)} e^{k_e})w$$

where $w \in \mathrm{min}(M_n)$ and the sum of the k_e does not exceed $\ell(R)$. This set is finite and each element in it is equal to u modulo R, so we will find all of these elements eventually. Then let M_{n+1} stand for the union of these sets.

By Lemma 5.1 there is a number n such that

$$\min(M_n) = \min(M_{n+1}), \min(E_n) = \min(E_{n+1}).$$

We claim that then $M_{\min}$ is equal to $\min(M_n)$, $E_{\min}$=$\min(E_n)$. Indeed, it is enough to show that if we apply a relation from R to a word

$$v = (\prod_{e \in \min(E_n)} e^{k_e})w \tag{7}$$

where $w \in \min(M_n)$ then we obtain a word w of the same form. If the sum of the k_e does not exceed $\ell(R)$ then this follows from the definition of M_n and from the equalities $\min(M_n)$=$\min(M_{n+1})$, $\min(E_n)$=$\min(E_{n+1})$.

Let this sum be greater than $\ell(R)$. Any application of a relation from R touches at most $\ell(R)$ units e from the right hand part of (7). Therefore $v = (\prod_{e \in \min(E_n)} e^{m_e})v_1$, and $w = (\prod_{e \in \min(E_n)} e^{m_e})w_1$ where v_1 belongs to M_n and w_1 is obtained from v_1 by applying a relation from R. But then w_1 is of the form (7), and so is w. This completes the proof.

5.3.2. *Nilpotent-by-Abelian Groups*
Let us consider the group variety $\mathcal{N}_2\mathcal{A}$ and prove that the word problem is decidable there. The proof we give essentially belongs to Kharlampovich [44]. See also the survey [45]

Let us take any group $G = < X >$, finitely presented in $\mathcal{N}_2\mathcal{A}$. The group G may be represented as a factor of the relatively free group $\bar{F} = < X >$ of $\mathcal{N}_2\mathcal{A}$.

Therefore $G = \bar{F}/R$ for some normal subgroup R which is finitely generated as a normal subgroup. Given a word w over X, we want to decide if $w = 1$ in G. We can consider w as an element of $\bar{F}$. Then $w = 1$ in G if and only if w belongs to R as an element of $\bar{F}$. Thus the word problem in G is decidable if and only if the membership problem for R in $\bar{F}$ is decidable.

We can find a somewhat bigger normal subgroup for which the membership problem *is* decidable. Consider the subgroup $\bar{F}''R$ (here $\bar{F}''$ is the second derived subgroup of $\bar{F}$). The factor group $\bar{F}/\bar{F}''R$ is finitely presented in the variety of metabelian groups. It is known that the word problem in the variety of metabelian groups is decidable [37]. Therefore the membership problem for the subgroup $\bar{F}''R$ is decidable.

So we can check if our word w belongs to $\bar{F}''R$. If $w \notin \bar{F}''R$ then $w \notin R$. So suppose that $w \in \bar{F}''R$. Since $w \in \bar{F}''R$, we can represent w as a product pr with $p \in \bar{F}''$, $r \in R$. This may be done effectively. Since $r \in R$ it is enough to decide if $p \in R$. The word p belongs to $\bar{F}''$. Thus we may suppose that $w \in \bar{F}''$. Now we have to consider the membership problem for the subgroup $R \cap \bar{F}''$.

Consider $\bar{F}''$ first. This is an Abelian subgroup. If $\bar{F}''$ were finitely generated then $R \cap \bar{F}''$ would be a finitely generated Abelian subgroup and the membership problem for $R \cap \bar{F}''$ would be trivially decidable. Unfortunately $\bar{F}''$ is not finitely generated as an Abelian subgroup (if $\bar{F}$ is not a cyclic group).

The idea is to define a module structure on $\bar{F}''$ to make $\bar{F}''$ a finitely generated module over a Noetherian commutative ring, because such modules have almost as nice algorithmic properties as finitely generated Abelian groups. There is a standard way to make $\bar{F}''$ a module.

Any Abelian normal subgroup A in a group H may be viewed as a right module over the group algebra $\mathbf{Z}H$ where the action of $\mathbf{Z}H$ on A is defined by the following formula:

$$a \circ (\sum \pm g_i) = \Pi a^{g_i^{\pm 1}}.$$

It is clear that this module has a big annihilator. For example $\mathbf{Z}A$ and even $\mathbf{Z}C_H(A)$ (here $C_H(A)$ is the centralizer of A in H) is in this annihilator. Therefore A may be considered as a module over $\mathbf{Z}H/N$ for every normal subgroup $N \leq C_H(A)$.

In our case $\bar{F} \in \mathcal{N}_2\mathcal{A}$. Therefore $\bar{F}''$ is Abelian and $C_{\bar{F}}(\bar{F}'') \geq \bar{F}'$. Hence $\bar{F}''$ may be considered as a right module over the ring $K = \mathbf{Z}\bar{F}/\bar{F}'$. This ring is just the ring of polynomials over Z with $X \cup X^{-1}$ as the set of unknowns factored by the ideal generated by elements $xx^{-1} - 1$ for all $x \in X$. Thus this is a commutative, finitely generated Noetherian domain.

Unfortunately $\bar{F}''$ is not finitely generated as a module over K. But we can *make* it a finitely generated module by increasing K! Indeed, by definition, the second derived subgroup $\bar{F}''$ consists of all elements which are products of double commutators $[[a,b],[c,d]]$ where $a,b,c,d \in \bar{F}$. Let us define a new action of generators from X on $\bar{F}''$:

$$[[a,b],[c,d]] \circ x = [[a,b]^x,[c,d]][[a,b],[c,d]^x].$$

This action is well defined and can be extended to another action of $K = Z\bar{F}/\bar{F}'$ on F''. Moreover this action commutes with the first action of K. Thus we can consider $\bar{F}''$ as a module over the tensor product $K \otimes K$. The last ring is, of course, also a finitely generated domain. It can be verified that $\bar{F}''$ is a finitely generated module over $K \otimes K$. Thus *we have found a hidden module structure on* $\bar{F}''$.

But we are mainly interested in R, or more precisely, in $R \cap \bar{F}''$. This is a normal subgroup and thus a K-submodule under the first action of K (normal subgroups are closed under conjugations). But this is not necessarily a $K \otimes K$-submodule. Forunately we can split $(R \cap \bar{F}'')$ into a a sum of three Abelian groups $A_1 + A_2 + A_3$ where A_1 is finitely generated as a subgroup ($\mathbf{Z}$-module), A_2 is finitely generated as a normal subgroup (K-module under the first action of K) and A_3 is finitely generated as a $K \otimes K$-module. Moreover generators of A_1, A_2, A_3 may be found effectively (see [44] for details). Therefore the membership problem for R has been reduced to the membership problem for the sum of finitely generated modules A_1, A_2, A_3 over different rings.

Now we can apply the following powerful result by Romanovskii.

Lemma 5.2 *[97] Let M be a $D-$module where D is a finitely generated commutative domain. Let $D_1 < D_2 < \ldots < D_m < D$ be a sequence of finitely generated subrings of D. Let N_i $(i = 1, \ldots, m)$ be a finitely generated D_i-submodule of M. Then the membership problem for $N = N_1 + N_2 + \ldots + N_m$ is decidable.*

This lemma gives us the decidability of the membership problem for $R \cap \bar{F}''$, which, in turn, implies the decidability of the word problem in $\mathcal{N}_2\mathcal{A}$.

5.3.3. *Other Hidden Structures*
It is interesting that a hidden module structure appears in the case of semigroups also. We will call a semigroup S a *semi-module* over a semigroup A if there exists an action $\circ$ of A on S (i.e. a function $\circ : A \times S \to S$ with $(a_1a_2) \circ s = a_1 \circ (a_2 \circ s)$) which "agrees" with the operation in S: $a \circ (s_1s_2) = (a \circ s_1)s_2 = s_1(a \circ s_2)$ for every $a \in A, s_1, s_2 \in S$. The following result was proved in [104].

Theorem 5.8 *[104] Let T be a semigroup finitely presented in a variety $\mathcal{V}$ satisfying one of the following identities*[3]

$$x^{nk}y(z^kt^k)^pz^{(n-p)k} = x^{(n-p)k}(t^kx^k)^pyz^{nk};$$
$$xy^nz = y^kxy^mzy^p, n > m.$$

and let every periodic group in $\mathcal{V}$ be locally finite. Then there exists a semi-module S over a commutative semigroup A such that T is "almost" a subsemigroup of S. More precisely there is a semigroup S_1 such that S is an ideal of S_1 with S_1/S finite and nilpotent and T is a factor-semigroup of S_1 over a congruence with all congruence classes finite.

This theorem is one of the main steps in describing semigroup varieties with decidable word problem.

The module structure is not the only possible hidden structure on the set of words equal to a given word. For example if we are dealing with inverse semigroups, it is useful to consider the Schützenberger automata discussed in Section 2.3. Recall that these automata accept all words which are greater than or equal to the given word. Then these automata may have nice geometric properties as in [68], which helps to solve the word problem (or to prove its undecidability). Another possiblity was discovered by Margolis and Meakin in [55]. They considered inverse semigroups S given by finitely many defining relations of the form $e = f$ where e and f are words which are idempotents in the free inverse semigroup (equivalently which are equal to 1 in the free group). In this case the Schützeberger graphs embed into the Cayley graph of the free group which is, of course, a labeled tree. Then the set of vertices of the Schützenberger graph may be defined by a formula from a decidable fragment of the second order theory of this tree. This implies the decidability of the word problem in S.

6. Minsky Machines and the Undecidability of the Word Problem in Varieties

6.1. Minsky Machines

One of the most powerful tools in proving the undecidability of the word problem is the so-called Minsky Machine. It was invented by Marvin Minsky in 1961 (see [72], [73], [53]). Yu. Gurevich was the first to use Minsky machines to prove the undecidability of an algorithmic problem in algebra [35].

[3] As it was mentioned earlier in this section, this condition is necessary for the decidability of the word problem.

Let us give a definition of a Minsky machine. The hardware of a (two-tape) Minsky machine consists of two tapes and a head. The tapes are infinite to the right and are divided into infinitely many cells numbered from the left to the right, starting with 0. The first cells on both tapes always contain 1, all other cells have 0. The head may acquire one of several internal states: $q_0, \dots, q_N$; q_0 is called *the terminal state*. At every moment the head looks at one cell of the first tape and at one cell of the second tape. So the *configuration* of the Minsky machine may be described by the triple (m, q_k, n) where m (resp. n) is the number of the cell observed by the head on the first (resp. second) tape, q_i is the state of the head.

Every command has the following form:

$$q_i, \epsilon, \delta \longrightarrow q_j, T^{\alpha}, T^{\beta}.$$

where $\epsilon, \delta \in \{0,1\}, \alpha, \beta \in \{-1,0,1\}$. This means that if the head is in the state q_i and it observes a cell containing ϵ on the first tape and a cell containing δ on the second tape then it acquires the state q_j and the first (the second) tape is shifted α (resp. β) cells to the left relative to the head. If, say, $\alpha = -1$ then the first tape is shifted one cell to the right. The machine always starts working at state q_1 and ends at the *stop state* q_0. The program (software) for a Minsky machine is a set of commands of the above form.

One says that a Minsky machine *calculates* a function $f(m)$ if for every m starting at the configuration $(m, q_1, 0)$ it ends at the configuration $(f(m), q_0, 0)$. If m does not belong to the domain of f then the machine works forever and never gets to the terminal state.

The main property of Minsky machines is contained in the following theorem.

Theorem 6.1 *[72] For every partially recursive function $f(m)$ there exists a Minsky machine which calculates the partial function $g_f : 2^m \to 2^{f(m)}$.*

Remark. It is interesting that, in this theorem, the function g_f cannot be replaced by the function f. In particular, there is no Minsky machine which calculates the function m^2 (see [53]).

The "canonical" definition of a Minsky machine seems to be very long and complex while in fact it is very simple and can be understood by a high school or even elementary school student. Let us give a "high school" definition of a Minsky machine.

Consider two glasses. We assume that these glasses are of infinite height. Another (more restrictive!) assumption is that we have infinitely many coins. There are four *operations*: "Put a coin in a glass", "Take a coin from a glass if it is not empty". We are able to check if a glass is/is not empty. A *program* is a numbered sequence of instructions.

An *instruction* has one of the following forms:

- Put a coin in the glass # n and go to instruction # j;
- If the glass # n is not empty then take a coin from this glass and go to instruction # j otherwise go to instruction # k.
- Stop.

A program starts working with the command number 1 and ends when it comes to the Stop instruction which will always have number 0.

We say that a program calculates a function $f(m)$ if, starting with m coins in the first glass and empty second glass, we end up with $f(m)$ coins in the first glass and empty second glass.

This "high school" version of a Minsky machine is known also as a Minsky algorithm. One can prove that Minsky machines are equivalent to Minsky algorithms, that is, given a program for Minsky machine (resp. given a Minsky algorithm), it is easy to construct a Minsky algorithm (resp. a program for Minsky machine) which calculates the same function.

A configuration of a Minsky algorithm is a triple (m, k, n), where m is the number of coins in the first glass, n is the number of coins in the second glass, and k is the number of the instruction we are executing. So the number of an instruction in the algorithm plays the role of an inner state.

Notice that although Minsky machines and Minsky algorithms are equivalent, Minsky algorithms are sometimes better tools in proving the undecidability of algorithmic problems. The main reason: there are four possible commands of a Minsky machine which correspond to the same inner state q_k, while there is only one instruction of a Minsky algorithm with a given number k.

6.2. Minsky Machines and the Word Problem for General Algebras

Here we will show how to apply Minsky machines (algorithms) to prove the undecidability of the word problem. All applications of Minsky machines (algorithms) are based on the following idea. We will show this idea for Minsky machines; exactly the same idea works for Minsky algorithms. Take a Minsky machine M calculating a function g_f with a non-recursive domain X.

One can define an "equality" on the set of all possible configurations of the Minsky machine M: two configurations (m, q_k, n) and $(m', q_{k'}, n')$ are "equal" or "equivalent" (we will write $(m, q_k, n) \equiv (m', q_{k'}, n')$) if there exists another configuration $(m'', q_{k''}, n'')$ such that M transforms both configurations (m, q_k, n) and $(m', q_{k'}, n')$ into $(m'', q_{k''}, n'')$. It is easy to see that this "equality" is symmetric, transitive and reflexive. By the choice of M, $(m, q_1, 0) \equiv (1, q_0, 0)$ if and only if $m \in X$. Therefore we have a sequence of configurations γ_m and a special configuration γ_0 such that $\gamma_m \equiv \gamma_0$ iff $m \in X$.

Suppose now that we want to construct a finitely presented universal algebra $A(M)$ with an undecidable word problem. The idea of doing this has its origin in the works of Markov and Post [61], [91]. First, with every configuration ψ one associates a word $w(\psi)$. This word is usually called a *canonical word*. Then with every command κ of the Minsky machine M one associates a finite set of defining relations R_κ. The algebra $A(M)$ will be defined by the relations from the union R of all R_κ (which is finite since we have only a finite number of commands) and usually some other relations Q which are in a sense "independent" of R. We need Q, for example, to make $A(M)$ satisfy a particular identity.

The algebra $A(M)$ will have an undecidable word problem if the following property holds:

$$\psi_1 \equiv \psi_2 \text{ if and only if } w(\psi_1) = w(\psi_2) \text{ in } A. \tag{8}$$

Indeed, in this case one cannot algorithmically decide if $w(m, q_1, 0) = w(1, q_0, 0)$ for the given number m.

We will say that we have an *interpretation* of the Minsky machine M in the algebra $A(M)$ if we have an assignment $\psi \to w(\psi)$ with the property (8).

Usually, in order to prove the property (8) one has to prove two lemmas.

Lemma 6.1 *If we can proceed from configuration ψ_1 to configuration ψ_2 using command κ, then we can proceed from the word $w(\psi_1)$ to the word $w(\psi_2)$ using relations from R_κ.*

Lemma 6.2 *If we can proceed from the word $w(\psi_1)$ to the word $w(\psi_2)$ by using relations from the union of all R_κ, then $\psi_1 \equiv \psi_2$.*

It is easy to see that Lemmas 6.1 and 6.2 imply property (8). Lemma 6.2 in most cases is more difficult to prove than Lemma 6.1. It is worth mentioning also that in order to prove Lemmas 6.1 and 6.2 we usually do not need any information about the function that is calculated by M.

Let us also make a remark about the case where we are constructing an algebra with undecidable word problem which is finitely presented in a variety $\mathcal{V}$. In this case we are allowed to use identities of $\mathcal{V}$ when we deduce relations of $A(M)$. Notice that unlike the relations of R_κ corresponding to commands of M, the relations obtained from identities of $\mathcal{V}$ have no connections with the Minsky machine, and can spoil the canonical words. So we have to make the canonical words resistant to applications of identities of $\mathcal{V}$. In the best case, the identities of $\mathcal{V}$ are not applicable to canonical words.

The procedure for constructing an algebra finitely presented in a variety $\mathcal{V}$ with undecidable word problem is roughly the following. First we temporarily forget about $\mathcal{V}$ and construct an interpretation of a Minsky machine M in an “absolutely” finitely presented algebra $A(M)$. We prove Lemma 6.1 for $A(M)$. Then we consider the factor algebra $\hat{A}(M)$ of $A(M)$ by the verbal congruence corresponding to $\mathcal{V}$, that is we identify all pairs of terms in $A(M)$ which are identically equal in $\mathcal{V}$. The algebra $\hat{A}(M)$ is finitely presented in $\mathcal{V}$. Then we have to prove Lemmas 6.1 and 6.2 for $\hat{A}(M)$. Fortunately we have Lemma 6.1 for free because the statement of this lemma is stable under homomorphic images. To prove Lemma 6.2 we usually need the above mentioned independence of canonical words from the identities of $\mathcal{V}$.

6.3. Why Minsky Machines?

The concept of an interpretation of a machine in an algebra can be applied to any kind of Turing machine. Why did we choose the Minsky machines? The first answer is: because Minsky machines are in some important sense the simplest universal Turing machines possible.

Indeed, the first and the main step in any interpretation of a machine in an algebra is the choice of canonical words. The canonical words must encode configurations of the machine. Therefore the smaller the number of parameters which determine the configurations, the more freedom we have in simulating the parameters. A configuration of a Minsky machine is determined just by three numbers: m, i, n. Here i runs over a finite set. Therefore one can encode each i by a separate letter q_i. It is also important that the commands of Minsky machines change those three numbers in a natural way. They add 0, 1 or -1 to m and n, and change i according to some simple rule. Therefore we can simulate m and n by, say, powers of different letters, say, a and b, and the relations corresponding to the relations of the Minsky machine will increase (decrease) the powers of these letters. Therefore we can encode the configuration (m, i, n) by a canonical word $a^m q_i b^n$ (we suppose for simplicity that we have a binary associative operation).

True, after a little pondering one can conclude that something is missing in this encoding. Indeed, recall that the relations will simulate the commands. The action of a command depends on whether m or n is equal to 0 or not, so there are four different situations ($m \neq 0, n \neq 0$; $m = 0, n \neq 0$; $m \neq 0, n \neq 0$; $m = 0, n = 0$). Therefore for each one of these four situations, the corresponding canonical word must have a special small subword which tells us that this situation occurs (then the corresponding relation will replace this subword by some other word).

Notice that the canonical words corresponding to these four situations have the following form: $a^n q_i b^m$, $q_i b^m$, $a^n q_i$, q_i where $m, n > 0$. All these words are subwords of the first one. So every subword of the second (the third or the fourth) word is a subword of the first word. Thus we cannot distinguish between these situations.

The solution to this problem is simple: we have to add two more letters A and B, which we will call "locks", and encode the configuration (m, i, n) by the word $Aa^m q_i b^n B$. Then each of the four situations is characterized by a small subword of the canonical word: The canonical word has a subword:

aq_ib iff $m \neq 0, n \neq 0$;

Aq_ib iff $m = 0, n \neq 0$;

aq_iB iff $m \neq 0, n = 0$;

Aq_iB iff $m = 0, n = 0$.

Thus it is very easy to find an interpretation of the Minsky machine. But this is not the only reason why one has to use it.

Recall that we want to simulate a machine in an algebra which satisfies as many identities as possible. And as everybody knows, those identities tend to change words. For example suppose that we have an identity $xy = yx$, and we simulate the configuration (m, i, n) by the word $Aa^m q_i b^n B$. Then the words $Aa^m q_i b^n B$, $a^m Aq_i b^n B$, $Aa^m q_i Bb^n$, and $a^m Aq_i Bb^n$ are equal and, again, we cannot find small subwords which distinguish one of the four situations from another.

The important feature of Minsky machines and their interpretations is that the canonical words, which we obtain, are very stable with respect to identities.

For example, in the case of semigroups, Theorem 5.4 means the following. There exist a few basic encodings of configurations of Minsky machines, and if a non-periodic semigroup variety with locally finite nil-semigroups satisfies identities which can change all these encodings, then the word problem is solvable in this variety, so the interpretation of a Minsky machine (and any other universal Turing machine) in this variety is impossible.

It is much more difficult to find stable interpretations of other kinds of machines.

For example, the configuration of the general (one-tape) Turing machine is a triple u, q_i, v where q_i is an inner state of the head, u is the word written on the tape to the left of the head, and v is the word written to the right of the head. So if we want to interpet the general Turing machine, we need to encode somehow arbitrary words u and v. Since any identity can change some words, we cannot encode the word by itself, thus the encoding must be unnatural. This will lead to great difficulties with simulating commands of the Turing machine, and so on.

6.4. The Word Problem for Semigroup Varieties. The Nonperiodic Case

There are two important semigroup interpretations of Minsky machines: the semigroups S_1 and S_2 below. Let M be a Minsky machine with internal states $q_0, \dots, q_N$. Then both S_1 and S_2 are generated by the elements $q_0, \dots, q_N$ and a, b, A, B. The correspondences between commands of M and relations of S_1 and S_2 are given by the following tables. Every command corresponds to one relation in S_1 and one relation in S_2.

Command	S_1
$q_i, 0, 0 \to q_j, T^\alpha, T^\beta$	$aq_ib = a^{1+\alpha}q_jb^{1+\beta}$
$q_i, 1, 0 \to q_j, T^\alpha, T^\beta$	$Aq_ib = Aa^\alpha q_jb^{1+\beta}$
$q_i, 0, 1 \to q_j, T^\alpha, T^\beta$	$aq_iB = a^{1+\alpha}q_jb^\beta B$
$q_i, 1, 1 \to q_j, T^\alpha, T^\beta$	$Aq_iB = Aa^\alpha q_jb^\beta B$

(9)

Command	S_2
$q_i, 0, 0 \to q_j, T^\alpha, T^\beta$	$q_iab = q_ja^{1+\alpha}b^{1+\beta}$
$q_i, 1, 0 \to q_j, T^\alpha, T^\beta$	$q_iAb = q_ja^\alpha Ab^{1+\beta}$
$q_i, 0, 1 \to q_j, T^\alpha, T^\beta$	$q_iaB = q_ja^{1+\alpha}b^\beta B$
$q_i, 1, 1 \to q_j, T^\alpha, T^\beta$	$q_iAB = q_ja^\alpha Ab^\beta B$

(10)

The canonical words in S_i are the following:

Configuration	S_1	S_2
(m, q_k, n)	$Aa^mq_kb^nB$	$q_ka^mAb^nB$

(11)

To make these interpretations work and to make these semigroups satisfy as many identities as possible we need also some additional relations independent of the commands of M.

In the semigroup S_2 we need the following commutativity relations:

$$ab = ba,\ aB = Ba,\ bA = Ab,\ AB = BA. \tag{12}$$

Also we need all relations of the type

$$xy = 0$$

where xy is a two letter word which is not a subword of $w(m, q_k, n)$ for some m, n or of any word obtained from $w(m, q_k, n)$ by the commutativity relations above.

These relations "kill" all "wrong" words. Basically only canonical words and their subwords are distinct from zero in S_1 and S_2.

Thus we have the following additional relations in S_1: all two letter words are equal to 0 except Aa, Aq_i, a^2, aq_i, q_ib, q_iB, b^2, bB. And we have the following additional relations in S_2: all two letter words are equal to 0 except q_ia, q_ib, q_iA, q_iB, a^2, aA, ab, aB, ba, b^2, bB, bA, Ab, AB, Ba, BA.

These semigroups are very convenient for demonstrating the standard proofs of Lemmas 6.1 and 6.2.

To prove Lemma 6.1 one needs to show that if we pass from a configuration (m, q_k, n) to another configuration $(m', q_{k'}, n')$ by a command κ then we pass from the word $w(m, q_k, n)$ to the word $w(m', q_{k'}, n')$ by the relation corresponding to κ.

Let us prove this only for the case of the semigroup S_2 and the command κ : $q_k, 1, 0 \to q_{k'}, T^\alpha, T^\beta$. All other cases are similar.

Since the command κ is applicable to the configuration (m, q_k, n), in this configuration, the head observes the first cell on the first tape and not the first cell on the second tape. Thus $m = 0, n \neq 0$. Then $m' = \alpha, n' = n + \beta$ (in this case α can not be negative). Now $w(m, q_k, n) = q_kb^nAB$ and the relation corresponding to κ is $q_kbA = q_{k'}a^\alpha b^{1+\beta}A$. Since $AB = BA$ and $aB = Ba$ we have $w(m, q_k, n) = q_kb^nAB = q_kbAb^{n-1}B$. Thus we can apply our relation and replace q_ibA by $q_{k'}a^\alpha b^{1+\beta}A$. As a result we obtain the word $q_{k'}a^\alpha b^{1+\beta}Bb^{n-1}A$ which is equal to $q_{k'}a^\alpha Ab^{n+\beta}B$ since $bA = Ab$. The last word is equal to $w(m', q_{k'}, n')$ as desired.

The proof of Lemma 6.2 is based on the following two standard observations. First, for every canonical word $w(m, q_k, n)$ there exists at most one relation corresponding to a command of M which is applicable to this word from the left to the right (this means that one replaces the left hand side of this relation by the right hand side of it).

Second, any application of a relation from tables (9) or (10) to any canonical word — from the left to the right or from the right to the left — gives us another canonical word (we do not distinguish words in S_2 which are obtained from each other by the commutativity relations[4]).

[4] Actually — important! — we assign to each configuration not a single word $w(m, q_k, n)$ but a set of words which may be obtained from each other by the commutativity relations. It is worth mentioning that it is almost always better to consider $w(m, q_k, n)$ as a set of words; relations corresponding to commands of M connect these sets of words; auxiliary relations connect words in the same set.

Now consider two words $w(m, q_k, n)$ and $w(m', q_{k'}, n')$ in S_1 or S_2. Suppose these words are equal in this semigroup. Therefore there exists a sequence of words

$$w(m, q_k, n), w_1, \ldots, w_n, w(m', q_{k'}, n')$$

where each word is obtained from the previous one by applying a defining relation corresponding to a command of the machine M. By the second observation each w_i corresponds to a configuration of M.

These relations may be applied from the left to the right and from the right to the left. Now suppose that in the passage $w_{r-1} \to w_r$, a relation was applied from the right to the left and in the passage $w_r \to w_{r+1}$, a relation was applied from the left to the right. Then w_{r-1} and w_{r+1} are obtained from w_r by applying relations from the left to the right. By the first observation these two relations must coincide and the words w_{r-1} and w_{r+1} must coincide also. Therefore in this case we can shorten our sequence of w_i. Thus we can suppose that in our sequence, there is a word $w_r = w(m'', q_{k''}, n'')$ such that all relations before w_r are applied from the left to the right and all relations after w_r are applied from the right to the left. But this means that the machine M passes from both configurations (m, q_k, n) and $(m', q_{k'}, n')$ to the configuration $(m'', q_{k''}, n'')$. Therefore $(m, q_k, n) \equiv (m', q_{k'}, n')$, as desired.

Thus we have proved that if the Minsky machine computes a non-recursive function then S_1 and S_2 have undecidable word problems. These semigroups are important because every (finitely based) semigroup variety with undecidable word problem, whose periodic semigroups are locally finite, contains either S_1 or S_2 or the semigroup anti-isomorphic to S_2 (see [104]). Therefore we do not need any other semigroup with undecidable word problem to treat non-periodic varieties with good periodic semigroups. But if the periodic semigroups are not locally finite we need something else, and we will discuss it in the next subsection.

6.5. The Word Problem for Semigroup Varieties. The Periodic Case

Theorem 4.5 above implies that a finitely based variety of semigroups in which not every periodic semigroup is locally finite contains either a periodic group or a nil-semigroup which is not locally finite. The periodic group case will be considered further (see Section 6.6). Now let us consider the nil-case. So suppose that we have a finitely based variety $\mathcal{V}$ containing a non-locally finite nil-semigroup. Then by virtue of Theorem 5.5, $\mathcal{V}$ has an undecidable word problem and now we are going to explain how to prove this using Minsky machines.

To show only the principal details of the proof let us prove that the word problem is undecidable in the variety of semigroups given by the identity $x^3 = 0$. This variety consists of nil-semigroups and it was proved by Morse and Hedlund [74] that it contains an infinite finitely generated semigroup. We will need a slight modification of the Morse and Hedlund construction.

Let us start with the following Thue endomorphism ϕ [122] of the free semigroup with generators a, b:

$$\phi(a) = ab, \ \phi(b) = ba.$$

Now let us iterate ϕ and consider the words $a, \phi(a), \phi^2(a), \phi^3(a), \ldots$. For every n, $\phi^n(a)$ is an initial segment of the word $\phi^{n+1}(a)$, so all these words are initial segments of an infinite sequence. Let us denote this sequence by $T(\phi)$. Thue [122] proved that these words do not have subwords of the form uuu where u is any non-empty word. Given $T(\phi)$, Morse and Hedlund construct a semigroup $S(\phi)$ as follows. Let $S(\phi)$ be the set consisting of all subwords of the words from this sequence, and a special symbol 0. Define an operation on this set by the following rule:

$$u \cdot v = \begin{cases} uv & \text{if } uv \text{ is a subword of } \phi^n(a) \text{ for some } n; \\ 0 & \text{if } uv \text{ is not a subword of any } \phi^n(a). \end{cases}$$

Here uv is the result of concatenation of u and v. [5]

It is easy to prove that $S(\phi)$ is an infinite semigroup generated by a and b. This semigroup satisfies the identity $x^3 = 0$. Indeed, as we have mentioned above, for every word u the word u^3 is not a subword of any $\phi^n(a)$ and so it is equal to 0 in $S(\phi)$.

One can easily see that it is not possible to use the interpretations from the previous subsection to simulate a Minsky machine in a semigroup satisfying $x^3 = 0$. Indeed, we can no longer encode the number of the cell observed by the head of the machine by a power of a letter: there is a shortage of powers (only 3). The idea is to use powers of the Thue endomorphism ϕ instead of powers of letters.

Thus we want to encode the configuration (m, q_k, n) by a word like $\phi^m(a)q_k\phi^n(a)$. Notice that such a word will be cube-free (will not contain subwords of the form uuu) for any m and n, so we won't be able to apply our identity $x^3 = 0$ to this word.

Now we have to assign a relation to every command of M. This relation must increase (decrease) m and n in $w(m, q_k, n)$ if the command shifts the tapes to the left (to the right). Unfortunately it is impossible to pass from $\phi^m(a)q_k\phi^n(a)$ to $\phi^{m+1}(a)q_{k'}\phi^n(a)$ by using one relation independent of m and n. Indeed, for "large" m it is impossible to proceed from $\phi^m(a)$ to $\phi^{m+1}(a)$ by replacing a "small" subword by another "small" subword.

But we notice that $\phi^{m+1}(a) = \phi(\phi^m(a))$, so we need to find some auxiliary relations which will simulate the application of ϕ. This can be done by adding one letter say, c_1 and relations $ac_1 = c_1\phi(a), bc_1 = c_1\phi(b)$. Indeed, for every m we will then have $\phi^m(a)c_1 = c_1\phi^{m+1}(a)$.

A practical realization of this idea is the following (see Section 6.6 for another realization).

Our semigroup, — let us denote it by $S(M, \phi)$,— will be generated by the set $\{q_0, \ldots, q_N,\ a, b,\ c_1, c_0, c_{-1},\ d_1, d_0, d_{-1},\ A, B\}$.

For every configuration (m, q_k, n) let

$$w(m, q_k, n) = A\overline{\phi^m(a)}c_0q_kd_0\phi^n(a)B.$$

where $\bar{u}$ is the word u written from the right to the left.

The correspondence between commands of the Minsky machine M and relations in $S(M, \phi)$ is given by the following table.

[5] As we mentioned in Section 4 there is a general construction which associates a semigroup $S(D)$ with every symbolic dynamical system D. The semigroup $S(\phi)$ is equal to $S(D)$ where D is the symbolic dynamical system generated by the limit of words $\phi^n(a)$.

Command	$S(M,\phi)$
$q_i, 0, 0 \to q_j, T^\alpha, T^\beta$	$bac_0q_id_0ab = bac_\alpha q_jd_\beta ab$
$q_i, 1, 0 \to q_j, T^\alpha, T^\beta$	$Aac_0q_id_0ab = Aac_\alpha q_jd_\beta ab$
$q_i, 0, 1 \to q_j, T^\alpha, T^\beta$	$bac_0q_id_0aB = bac_\alpha q_jc_\beta aB$
$q_i, 1, 1 \to q_j, T^\alpha, T^\beta$	$Aac_0q_kd_0aB = Aac_\alpha q_jd_\beta aB$

(13)

The auxiliary relations are the following:

(i) $ac_0 = c_0a$, $bc_0 = c_0b$;

(ii) $ac_1 = c_1\phi(a)$, $bc_1 = c_1\phi(b)$, $Ac_1 = Ac_0$;

(iii) $\phi(a)c_{-1} = c_{-1}a$, $\phi(b)c_{-1} = c_{-1}b$, $Ac_{-1} = Ac_0$;

(iv) $d_0a = ad_0$, $d_0b = bd_0$;

(v) $d_1a = \phi(a)d_1$, $d_1b = \phi(b)d_1$, $d_1B = d_0B$;

(vi) $d_{-1}\phi(a) = ad_{-1}$, $d_{-1}\phi(b) = bd_{-1}$, $d_{-1}B = d_0B$.

The role of the new generators c_i and d_i is clear from these relations: d_1 and c_1 increase the power of ϕ, c_{-1} and d_{-1} decrease the power of ϕ.

It is not very difficult to prove Lemmas 6.1 and 6.2. Therefore we have obtained an interpretation of the machine M. Now, since all words $w(m, q_k, n)$ and all words which can be obtained from these words by the defining relations of $S(M,\phi)$ are cube-free, the identities of our variety won't change these words. Therefore the statements of Lemmas 6.1 and 6.2 hold for the factor-semigroup $\hat{S}(M,\phi)$ of the semigroup $S(M,\phi)$ over the verbal congruence corresponding to the identity $x^3 = 0$. Indeed, the statement of Lemma 6.1 is stable under homomorphic images. The statement of Lemma 6.2 holds because canonical words which are distinct in $S(M,\phi)$ are also distinct in $\hat{S}(M,\phi)$. It remains to notice that the semigroup $\hat{S}(M,\phi)$ is finitely presented in the variety given by the identity $x^3 = 0$. Therefore this variety has an undecidable word problem.

In the general case, when the identities of the variety $\mathcal{V}$ are more (sometimes much more) complicated, one has to use the endomorphisms constructed in [108] instead of ϕ, and the interpretation is slightly different also (see [105]).

6.6. The Word Problem for Varieties of Groups. The Periodic Case

Here we present the method of Sapir from [106] of constructing a group with undecidable word problem which is finitely presented in the variety $\mathcal{A}_r\mathcal{B}_p$ for every odd $p \geq 665$ and every prime $r \neq p$.

The free group in the variety $\mathcal{A}_r\mathcal{B}_p$ is a semidirect product of an Abelian group A of exponent r and the free Burnside group F of exponent p. This Abelian group may be considered as a module over the group ring $\mathbf{Z}_rF$. Standard group theoretic observations show that in order to construct a group which is finitely presented in this variety and has an undecidable word problem, it is enough to construct a finitely generated right ideal in the group ring $\mathbf{Z}_r B$ with undecidable membership problem, where B is a finitely presented group from $\mathcal{B}_p$.

In order to find this ideal we do the following. First of all we find a semigroup interpretation of a Minsky algorithm, similar to the interpretation given in the previous section, which satisfies the following properties.

1. Every canonical word is cube-free.
2. Every relation corresponding to a command of the Minsky algorithm applies to a prefix of the canonical word.
3. The auxiliary relations are some commutativity relations like $ab = ba, Ab = bA$, etc. and do not contain letters q_i which correspond to the numbers of commands of the algorithm.

The interpretation used in [106] is the following.

The alphabet consists of letters

$$\{q_i, q_{i1}, a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2, A, B, C, D \mid i = 0, 1, \dots, N\}.$$

Recall that a Minsky algorithm deals with coins in two glasses and can add (take off) a coin to (from) one of these glasses (see Section 6.1 for the definition of Minsky algorithms). A configuration of the algorithm is a triple (m, i, n) where i is the number of the command which is being executed, m is the number of coins in the first glass, n is the number of coins in the second glass.

Let ϕ be the Thue-like substitution (see Section 6.5):

$$\phi(x_1) = x_1 x_2, \ \phi(x_2) = x_2 x_1, \ x \in \{a, b, c, d\}.$$

Then the canonical word $w(m, i, n)$ is equal to $q_i \phi^m(a_1)\phi^n(b_1)ABCD$.

Auxiliary relations are the following:

$$x_i y_j = y_i x_j, \quad x_i Y = Y x_i \tag{14}$$

where $x, y \in \{a, b, c, d\}, x \neq y, Y$ is the capital y for $y \in \{a, \ b, \ c, \ d\}$.

These relations make letters $a, \ b, \ c, \ d$ "independent": they can move inside the canonical words without disturbing each other.

The relations corresponding to the commands of the Minsky algorithm are the following.

Command	Semigroup relations
[add a coin to the first glass | j]	$q_ia_t = q_i\phi(c_t), t = 1,2$ $q_iA = q_{i1}A$ $q_{i1}c_t = q_{i1}a_t, t = 1,2$ $q_{i1}C = q_jC$
[add a coin to the second glass | j]	$q_ib_t = q_i\phi(d_t), t = 1,2$ $q_iB = q_{i1}B$ $q_{i1}d_t = q_{i1}b_t, t = 1,2$ $q_{i1}D = q_jD$
[take a coin from the first glass | $j\|k$]	$q_ia_1A = q_ka_1A$ $q_i\phi(a_t) = q_ic_t, i = 1,2$ $q_iA = q_{i1}A$ $q_{i1}c_t = q_{i1}a_t$ $q_{i1}C = q_jC$
[take a coin from the second glass |$j\|k$]	$q_ib_1B = q_kb_1B$ $q_i\phi(b_t) = q_id_t, i = 1,2$ $q_iB = q_{i1}B$ $q_{i1}d_t = q_{i1}b_t$ $q_{i1}D = q_jD$

(15)

Then we take the alphabet A of this interpretation and produce a new alphabet A_1 which consists of of four copies of letters from A. If x is a letter of A then A_1 contains letters x_{ij} where $i, j \in \{1,2\}$. We assume that new letters x_{ij} satisfy the same auxiliary commutativity relations as x. For example, since we had a relation $ab = ba$, we now have relations $a_{ij}b_{k\ell} = b_{k\ell}a_{ij}$ for all i, j, k, ℓ.

We replace every letter x in our interpretation of the Minsky algorithm (that is in the canonical words and in the relations) by the sum $x_{11} + x_{12} + x_{21} + x_{22}$ which we consider as an element of the group algebra $\mathbf{Z}B$ where B is the Burnside group generated by A_1 subject to the new auxiliary relations.

The polynomials, obtained from a word u over A by this substitution will be denoted by $\bar{u}$.

Now let $R(M) = \{u_i = v_i \mid i = 1,2,\ldots\}$ be the set of relations which simulates our Minsky algorithm. Consider the right ideal $I(M)$ in $\mathbf{Z}B$ generated by the polynomials $\bar{u}_i - \bar{v}_i$, $i = 1,2,\ldots$.

It is not difficult to prove that if two configurations (m,i,n) and (m',i',n') of the Minsky algorithm are equivalent then the polynomial $\bar{w}(m,i,n) - \bar{w}(m',i',n')$ belongs to $I(M)$. One has to use the properties 2 and 3 of our interpretation (this is why the interpretation presented in the previous section does not work in the present case).

Thus in order to prove that $I(M)$ has an undecidable membership problem, it is enough to show that if $\bar{w}(m,i,n) - \bar{w}(m',i',n')$ belongs to $I(M)$ then the configurations (m,i,n) and (m',i',n') of the Minsky algorithm are equivalent.

Suppose that $\bar{w}(m,i,n) - \bar{w}(m',i',n')$ belongs to $I(M)$. This means that

$$\bar{w}(m,i,n) - \bar{w}(m',i',n') = \sum(\bar{u}_i - \bar{v}_i)f_i \tag{16}$$

where $u_i = v_i \in R,\ f_i \in B$.

Both sides become sums of monomials. Now take a monomial U in the sum $\bar{w}(m,i,n)$. From the theory of Burnside groups (see Adian [2]) we know that cube-free words are not equal to each other, and are not equal to smaller subwords in the free Burnside group. Therefore this monomial cannot coincide with other monomials from $\bar{w}(m,i,n)$ and with monomials from $\bar{w}(m',i',n')$. Therefore this monomial must coincide with a monomial from the right side of (16) (we are in a group ring!).

Hence we have $U = u_i'f_i$ or $U = v_i'f_i$. Here w' (for w equal to u_i or v_i) is obtained from w by replacing some letters by their "brothers" — the same letters with different indices). Suppose that the first equality holds. Suppose that u_i' is an initial part of the word U, i.e. the equality $U = u_i'f_i$ is an equality in the free semigroup. Then we can take a monomial v_i' in the sum $\bar{v}_i$ and replace U by $U_1 = v_i'f_i$. The monomial U_1 must cancel with some other monomial in a sum $(\bar{u}_s - \bar{v}_s)f_s$. Again suppose that $U_1 = u_s'f_s$ in the free semigroup, and replace U_1 by $U_2 = v_j'f_s$. Finally we will hit all sums $(\bar{u}_s - \bar{v}_s)f_s$ in the right side of the equality (16) and so one of U_g will coincide with a monomial from the sum $\bar{w}(m',i',n')$. Now look at the sequence of monomials $U, U_1, U_2, \ldots, U_k$. If we identify letters and their "brothers" then U will coincide with $w(m,i,n)$, U_k will coincide with $w(m',i',n')$, and every step in this sequence of words is obtained by replacing one side of relations from R by another side of this relation. Therefore we have a semigroup deduction of the relation $w(m,i,n) = w(m',i',n')$. But we already know that if such a relation holds then corresponding configurations of the Minsky machine are equivalent.

This was only an idea of the proof. In order to make this proof work we have to understand what to do if U, the monomial we are dealing with, is equal to $u_i'f_i$ but u_i' is not an initial part of U. Of course, we still can take a monomial v_i' and replace U by $U_1 = v_i'f_i$ but first it may eventually lead to big powers of letters which is bad since we are working with a periodic group B. Second, the transformations $U \to U_1 \to U_2 \ldots$ will not simulate semigroup transformations.

In order to avoid these difficulties Sapir defines a set of words E over the alphabet of generators A_1 with the following properties.

E1. Every word in E is cube-free.

E2. If w is a canonical word then one and only one monomial in $\bar{w}$ belongs to E.

E3. If the word f_i is contained in E and the word $u_i'f_i$ is also contained in E, then there exists a unique v_i' such that $v_i'f_i$ belongs to E.[6]

E4. If the word f_i does not belong to E but $u_i'f_i$ belongs to E then there exists a unique monomial u_i'' in $\hat{u}_i$ which is distinct from u_i' and is such that $u_i''f_i$ belongs to E.

[6] Here $u_i = v_i$ is a relation from (10).

E5. If w is a positive word from E, $f \in E$, and $w = u_i' f$, then u_i' is an initial part of w.

Given this set E, we can describe our transformations $U \to U_1 \to U_2 \to \dots$ precisely. Indeed, suppose again that we have the equality (16). We know that there exists (unique) monomial U in $\bar{w}(m,i,n)$ which belongs to E. Since all monomials in the left part of (16) are cube-free, U is not equal to any monomial in the left side, and so it must be equal to some $u_i' f_i$ from the right side.

Now if f_i belongs to E, we choose a (unique by E3) word v_i' in $\bar{v}$, and replace U by $U_1 = v_i' f_i$. Since U is a positive word, u_i' is an initial part of U, and so U_1 is also a positive word. In this case we use all monomials from the sum $(\bar{u}_i - \bar{v}_i) f_i$ which belong to E. Let us call this transformation $U \to U_1$ an $R1$-transformation.

If f_i does not belong to E we choose a (unique by E4) monomial u_i'' in $\bar{u}_i$ which is distinct from u_i' and such that $U_1 = u_i'' f_i$ belongs to E. In this case we will call our transformation $U \to U_1$ an $R2$-transformation.

If U_1 belongs to the sum $\bar{w}(m', i', n')$, our process ends. If not, we can find another sum on the right side of (16) where it belongs. Then we can proceed from U_1 to U_2 by an $R1$- or $R2$-transformation.

Why does this process simulate the process of semigroup deductions? Notice that if all transformations in the sequence $U \to U_1 \to U_2 \to \dots$ are $R1-$transformations then all these words are positive and the sequence indeed simulates a deduction of the relation $w(m,i,n) = w(m',i',n')$.

In the case when some of these transformations are $R2$-transformations Sapir proves that there exists another (perhaps even shorter) sequence of $R1$-transformations only, which connects the same words.

6.7. The Word Problem for Varieties of Inverse Semigroups

6.7.1. *Inverse Semigroups with Abelian Covers*

Recall that inverse semigroups form a variety of algebras of type $< 2, 1 >$. While inverse semigroups appear at first sight to form a class of algebras intermediate between groups and semigroups, there are many aspects of the theory that are drastically different from their counterparts in group theory or semigroup theory. Even the theory of the free inverse semigroup [75], [112], [55], [98] offers many surprises and interesting unsolved problems. In particular, the free inverse semigroup is not finitely presented as a semigroup [111]. The structure of the free inverse semigroup $FIS(X)$ on X was described in Section 2 above.

Our concern in this section is with the variety of those inverse semigroups which possess an E-unitary cover over some Abelian group. The concept of E-unitary cover is one of the fundamental concepts in the theory of inverse semigroups. It has many equivalent definitions (see [88]). One of the definitions is the following. Let S be an inverse semigroup of partial injective transformations of a set X. We say that S has an E-unitary cover over a group G if G has a faithful representation by permutations of a set Y containing X such that every transformation of S is a restriction of some permutation of G.

For every variety of groups $\mathcal{V}$ the class of inverse semigroups

$$\hat{\mathcal{V}} = \{S | S \text{ has an E-unitary cover over some group in } \mathcal{V}\}$$

is a variety of inverse semigroups. The following result summarizes some of the work of Petrich and Reilly [89] and Pastijn [85].

Theorem 6.2 *Let $\mathcal{V}$ be a variety of groups. Then $\hat{\mathcal{V}}$ is defined by the identities* $\{u^2 = u \mid u = 1$ *is an identity in* $\mathcal{V}\}$.

The variety $\hat{\mathcal{V}}$ may also be characterized in a different way. Recall that if $\mathcal{U}$ and $\mathcal{V}$ are varieties of universal algebras then the *Mal'cev product* $\mathcal{U}\mathcal{V}$ is the variety generated by all algebras A for which there exists a congruence ρ on A such that all ρ-classes which are subalgebras are in $\mathcal{U}$ and $A/\rho \in \mathcal{V}$.

For varieties of inverse semigroups $\mathcal{U}$ and $\mathcal{V}$ we also denote by $\mathcal{U} * \mathcal{V}$ the variety generated by all semidirect products $S * T$ where $S \in \mathcal{U}$, $T \in \mathcal{V}$. It follows from the work of Tilson [121] and O'Carroll [80] and the fact that the variety of semilattices $\mathcal{S}$ is "local", that for every group variety $\mathcal{V}$ we have $\mathcal{S}\mathcal{V} = \mathcal{S} * \mathcal{V}$. From Petrich [88] (Theorem 12.9.11) we also have:

Theorem 6.3 *Let $\mathcal{V}$ be any variety of groups. Then*

$$\hat{\mathcal{V}} = \mathcal{S} * \mathcal{V} = \mathcal{S}\mathcal{V}.$$

The free semigroups in the variety $\hat{\mathcal{V}}$ have been described by Margolis and Meakin [54]. Let $\mathcal{V}$ be a variety of groups, X a nonempty set and $F_X(\mathcal{V})$ the relatively free X-generated group in $\mathcal{V}$. Let $\Gamma_X(\mathcal{V})$ denote the Cayley graph of $F_X(\mathcal{V})$ relative to the set of generators X. Define

$$\begin{aligned} M_X(\mathcal{V}) \; = \{(\Gamma, g) | \Gamma & \text{ is a finite connected subgraph of } \Gamma_X(\mathcal{V}) \\ & \text{containing 1 and } g \text{ as vertices}, \Gamma \neq \{1\}\} \end{aligned}$$

with multiplication

$$(\Gamma_1, g_1)(\Gamma_2, g_2) = (\Gamma_1 \cup g_1 \cdot \Gamma_2, g_1 g_2).$$

(Here $g_1 \cdot \Gamma_2$ denotes the natural action (left translation) of g_1 on Γ_2.)

The inverse monoid $M_X(\mathcal{V})^1 = M_X(\mathcal{V}) \cup \{1\}$ is just the inverse monoid $M(X : R)$ constructed in Section 2, corresponding to the natural presentation $gp < X : R >$ of $F_X(\mathcal{V})$.

From Margolis and Meakin ([54], Corollary 2.10) we have the following description of the relatively free semigroups in $\hat{\mathcal{V}}$

Theorem 6.4 *If $\mathcal{V}$ is any variety of groups and X is any non-empty set then $M_X(\mathcal{V})$ is the relatively free X-generated inverse semigroup in the variety $\hat{\mathcal{V}}$.*

In particular, if $\mathcal{V} = \mathcal{G}$, the variety of all groups, then $\hat{\mathcal{G}}$ is the variety of all inverse semigroups, and $M_X(\mathcal{G})$ is the free inverse semigroup on X.

Notice that from Theorem 6.4 it follows that the word problem in $M_X(\mathcal{V})$ is decidable if and only if the group $F_X(\mathcal{V})$ has decidable word problem.

If $\mathcal{A}$ is the variety of Abelian groups and $|X| = n$, then the Cayley graph $\Gamma_X(\mathcal{A})$ is the lattice $\mathbf{Z}^n$. The multiplication in $M_X(\mathcal{A})$ may be easily visualized. The word problem in $M_X(\mathcal{A})$ may be solved in the following manner. Let w be a word in $(X \cup X^{-1})^+$ and let Γ_w be the finite subgraph of the lattice $\Gamma_X(\mathcal{A})$ traversed when the word w is read as the label of a path in $\Gamma_X(\mathcal{A})$, starting at the vertex 1. Then $w_1 = w_2$ in $M_X(\mathcal{A})$ if and only if $\Gamma_{w_1} = \Gamma_{w_2}$ and the paths labelled by w_1 and w_2 in $\Gamma_X(\mathcal{A})$ have the same end (that is these words are equal in the free Abelian group). For example, if $X = \{a, b\}$, $w_1 = abaa^{-1}bb^{-1}a^{-1}b^{-1}b^2aa^{-2}b^{-1}a^2$, $w_2 = ba^{-1}ba^2a^{-1}b^{-1}abb^{-2}a^{-1}ba^2a^{-1}$ then $w_1 = w_2$ in $F_X(\mathcal{A})$ and $\Gamma_{w_1} = \Gamma_{w_2}$ is the graph shown in the Figure 2 below.

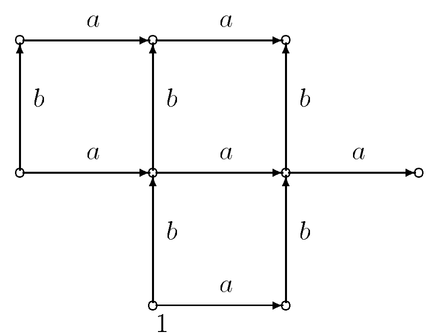

Fig. 2.

Hence $w_1 = w_2$ in $M_X(\mathcal{A})$. But it is easy to see that $w_1 \neq ab$ in $F_X(\mathcal{A})$, because $\Gamma_{ab} \neq \Gamma_{w_1}$.

Thus the word problem is easily decidable in the free semigroups of the variety $\hat{\mathcal{A}}$. In this section, however, we outline a proof of the fact that the word problem is undecidable in the variety $\hat{\mathcal{A}}$; that is we can add finitely many relations to the defining relations of $M_X(\mathcal{A})$ and obtain a semigroup with undecidable word problem. Full details of the proof of this result may be found in the paper [68].

More precisely we have the following

Theorem 6.5 *If $\mathcal{B}$ is any variety of inverse semigroups containing the variety $\hat{\mathcal{A}}$ of inverse semigroups possessing an E-unitary cover over an Abelian group, then $\mathcal{B}$ has undecidable word problem.*

The proof of this theorem uses the Schützenberger automata introduced in Section 2 associated with an inverse semigroup presentation and also an interpretation of a Minsky algorithm.

It is well-known that the variety $\hat{\mathcal{A}}$ contains all Brandt semigroups B_n and the bicyclic semigroup $B = Inv < a : aa^{-1} = 1 >$. The inverse semigroup $Inv < a, b : ab = ba >$ is in this variety but the inverse semigroup $Inv < a, b, c\ :\ ab = ba,\ ac = ca,\ bc = cb >$ is not because the Schützenberger automaton of the word $ab^{-1}ca^{-1}bc^{-1}$ is linear, so this word is not an idempotent in this semigroup, but it is a commutator and so it must be an idempotent in every semigroup of $\hat{\mathcal{A}}$ (see theorem 6.2).

6.7.2. *Minsky Algorithms and the Word Problem in Inverse Semigroups*
Let us take a Minsky algorithm M which calculates the characteristic function $f(n)$ of a recursively enumerable but non-recursive set X (i.e. $f(x) = 0$ if $x \in X$ and is not defined if $x \notin X$). We assume that M has $N + 1$ commands, $0, 1, \ldots, N$ (where 0 is the Stop command and 1 is the start command).

Recall that by the definition of an interpretation of a Minsky algorithm, to interpret a Minsky algorithm in a semigroup means that one must do the following:

- Associate with each configuration (m, i, n) of a Minsky algorithm, a canonical word $w(m, i, n)$;
- Associate with each command κ of the Minsky algorithm a set of defining relations R_κ.

These canonical words and sets of relations must have the following property:

$$(m, i, n) \equiv (m', i', n') \text{ iff } w(m, i, n) = w(m', i', n) \text{ modulo the union of all } R_\kappa.$$

Notice that the choice of the relations R_κ is in some sense determined by the choice of canonical words because these relations must simulate the execution of the commands of the Minsky algorithm.

The canonical word must encode the information contained in the configuration. The simplest words which encode the triple (m, i, n) are $Aa^m qb^n B$ and $qa^m Ab^n B$ (see Section 6.4)

In the case of inverse semigroups, these choices of canonical words do not suffice. Consider, for example the first choice. The Minsky algorithm M may certainly have equivalent configurations (m, i, n) and (m', i', n') with $i = i', n = n'$ but $m \neq m'$. In such a case the Schützenberger automaton of $w(m, i, n)$ must accept $w(m', i, n)$ and visa versa. For simplicity let $m = 2, m' = 4, n = 1$ (the general case is similar).

Thus this automaton must have paths as shown in Figure 3.

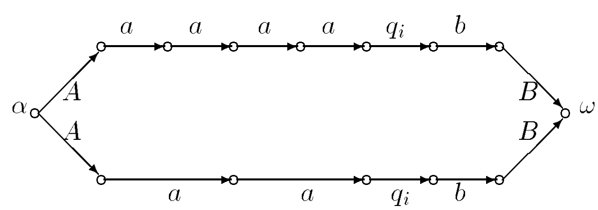

Fig. 3.

The requirement that the Schützenberger automaton of $w(m, i, n)$ be an inverse automaton (i.e. it must be both deterministic and injective), and the fact that $m \neq m'$ would then give rise to a subgraph of the Schützenberger automaton of the form shown in Figure 4.

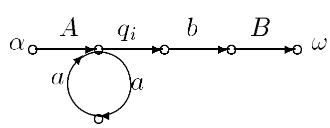

Fig. 4.

It follows that the Schützenberger automaton of $w(m,i,n)$ would accept words of the form $Aa^{2r}q_ib^nB$ for all r. This would force $(2r,i,n)\equiv(2,i,n)$ for every r.

In general it is easy to see that if $(m,i,n)\equiv(m+t,i,n)$ for some $m,n,t>0$ then $(m+tr,i,n)\equiv(m,i,n)$ for every r. From this, it is easy to deduce that our Minsky algorithm cannot compute the characteristic function of a non-recursive set X. Indeed, let Y be the set of all numbers m such that $(m,1,0)\equiv(1,0,0)$. We proved that with every two numbers $m, m+t$ the set Y contains all numbers of the form $m+rt$, $r=0,1,\ldots$. Let us take $m\in Y$ and $t>0$ such that $m+t\in Y$ and t is the smallest difference between elements of Y. Then it is clear that all elements from Y which are greater than m belong to the arithmetic progression $m, m+t, m+2t\ldots$. Hence Y contains only finitely many numbers outside this progression, so it is recursive. This contradicts the fact that by Minsky's theorem, 2^k belongs to Y if and only if k belongs to X. It is only slightly more difficult to prove that our algorithm cannot compute any non-recursive function: we leave this to the reader as an exercise. Difficulties of this kind lead to the consideration of longer and more symmetric canonical words, and hence to a longer list of defining relations than has been used for similar problems in varieties of semigroups and groups.

In [68] the following canonical word corresponding to the configuration (m,i,n) was chosen:

$$w(m,i,n)=A(i)a(i)^mB(i)b(i)^nQ(i)\bar b(i)^n\bar B(i)\bar a(i)^m\bar A(i).$$

The task then is to construct defining relations corresponding to each type of command of the Minsky algorithm M so that the relations simulate the execution of the algorithm. For example, in [68] the following set of defining relations corresponds to a command number i of the type

$$[\text{take a coin from the first glass } \mid j \mid k]. \tag{17}$$

(This means that if the first glass is not empty, then take a coin from it and pass to the command number j, and if the first glass is empty then pass to the command number k.)

1. $Q(i)=f(i)P(i)\bar f(i)$,

2. $b(i)f(i)=f(i)b_1(i)$, $\quad \bar f(i)\bar b(i)=\bar b_1(i)\bar f(i)$,

3. $a(i)B(i)f(i)=a(i)e(i)B_1(i)$, $\quad \bar f(i)\bar B(i)\bar a(i)=\bar B_1(i)\bar e(i)\bar a(i)$,

4. $a(i)e(i)=e(i)a_1(i)$, $\quad \bar e(i)\bar a(i)=\bar a_1(i)\bar e(i)$,

5. $A(i)e(i)=A(j)e_{3,1}(i)$, $\quad \bar e(i)\bar A(i)=\bar e_{3,1}(i)\bar A(j)$,

6. $e_{3,1}(i)a_1(i)=a(j)e_{3,1}(i)$, $\quad \bar a_1(i)\bar e_{3,1}(i)=\bar e_{3,1}(i)\bar a(j)$,

7. $e_{3,1}(i)a_1(i)B_1(i)=B(j)f_{3,1}(i)$, $\quad \bar B_1(i)\bar a_1(i)\bar e_{3,1}(i)=\bar f_{3,1}(i)\bar B(j)$,

8. $f_{3,1}(i)b_1(i)=b(j)f_{3,1}(i)$, $\quad \bar b_1(i)\bar f_{3,1}(i)=\bar f_{3,1}(i)\bar b(j)$,

9. $f_{3,1}(i)P(i)\bar{f}_{3,1}(i) = f_{3,1}(i)f_{3,1}^{-1}(i)Q(j)$

10. $A(i)B(i)f(i) = A(k)B(k)f_{3,2}(i), \qquad \bar{f}(i)\bar{B}(i)\bar{A}(i) = \bar{f}_{3,2}(i)\bar{B}(k)\bar{A}(k),$

11. $f_{3,2}(i)b_1(i) = b(k)f_{3,2}(i), \qquad \bar{b}_1(i)\bar{f}_{3,2}(i) = \bar{f}_{3,2}(i)\bar{b}(k),$

12. $f_{3,2}(i)P(i)\bar{f}_{3,2}(i) = f_{3,2}(i)f_{3,2}^{-1}(i)Q(k)$

Similar relations were introduced to simulate the execution of the other types of commands of the Minsky algorithm. The Stop command is simulated by the single defining relation $w(1,0,0) = 0$. This leads to a long list of defining relations and to an inverse semigroup $S(M)$ generated by the set $G(M)$ of all symbols involving in the defining relations and subject to the union of all of these relations.

Now let $\mathcal{B}$ be any variety of inverse semigroups containing $\hat{\mathcal{A}}$ and let $S_{\mathcal{B}}(M)$ be the inverse semigroup in $\mathcal{B}$ obtained by adding to the defining relations of $S(M)$ the identities of $\mathcal{B}$ (i.e. $S_{\mathcal{B}}(M)$ is defined in $\mathcal{B}$ by the same relations which define $S(M)$ in the class of all inverse semigroups).

We want to show that $w(2^m, 1, 0)$ is equal to 0 (=$w(1,0,0)$) in $S_B(M)$ if and only if $m \in X$. This is accomplished by proving the following two lemmas.

Lemma 6.3 *If we can pass from configuration (m,i,n) to configuration (m',j,n') using command i of M then $w(m,i,n) \leq w(m',j,n')$ in $S(M)$ where $\leq$ is the natural order on the inverse semigroup $S(M)$.*

Lemma 6.4 *If $m \notin X$ then $w(2^m,1,0)$ is not equal to 0 in $S_{\mathcal{B}}(M)$.*

Indeed, suppose that we have proved these two lemmas. If $m \in X$ then our algorithm M takes configuration $(2^m,1,0)$ to configuration $(1,0,0)$ in a finite number of steps. By Lemma 6.3 we can deduce that $w(2^m,1,0) \leq w(1,0,0) = 0$. But zero is the smallest element in the natural order of any inverse semigroup. Therefore $w(2^m,1,0) = 0$ in $S(M)$. Since $S_{\mathcal{B}}(M)$ is a homomorphic image of $S(M)$ we have that $w(2^m,1,0) = 0$ in $S_{\mathcal{B}}(M)$. On the other hand if $m \notin X$ then by Lemma 6.4 the element $w(2^m,1,0)$ is not equal to 0 in $S_{\mathcal{B}}(M)$. This will complete the proof of Theorem 6.5.

In order to prove lemma 6.3 it suffices to show that if we can proceed from configuration (m,i,n) to configuration (m',j,n') using command i of M then $w(m',j,n')$ must be accepted by the Schützenberger automaton $A(w(m,i,n))$ corresponding to this presentation. For example, suppose that command i is of the type 17. Start with the linear automaton of the word

$$w(m,i,n) = A(i)a(i)^m B(i)b(i)^n Q(i)\bar{b}(i)^n \bar{B}(i)\bar{a}(i)^m \bar{A}(i).$$

If $m \neq 0$ then application of command i to $w(m,i,n)$ must produce $w(m-1,j,n)$ while if $m = 0$, it must produce $w(m,k,n)$.

Notice that the relation 1 may be used to apply an expansion to the automaton (add a new path labelled by $f(i)P(i)\bar{f}(i)$ to the automaton starting at the initial vertex of the edge labelled by $Q(i)$ and ending at the terminal vertex of this edge). This then introduces a path labelled by $b(i)f(i)$ and a path labelled by $\bar{f}(i)\bar{b}(i)$, to which we can apply expansions by use of the relations 2 and so on. Continue in this manner: one easily checks that if one applies only the relations 1-12 corresponding to command i, then the automaton obtained from the linear automaton of $w(m,i,n)$ and closed with respect to application of expansions corresponding to these relations is in fact finite.

Sketches of these automata in the cases $m=3,\ n=3$ and $m=0,\ n=3$ are provided in figures 5 and 6.

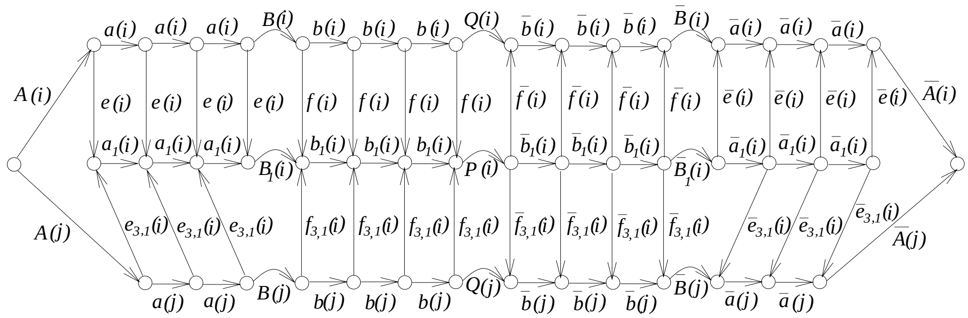

Fig. 5. $m=3,\ n=3$

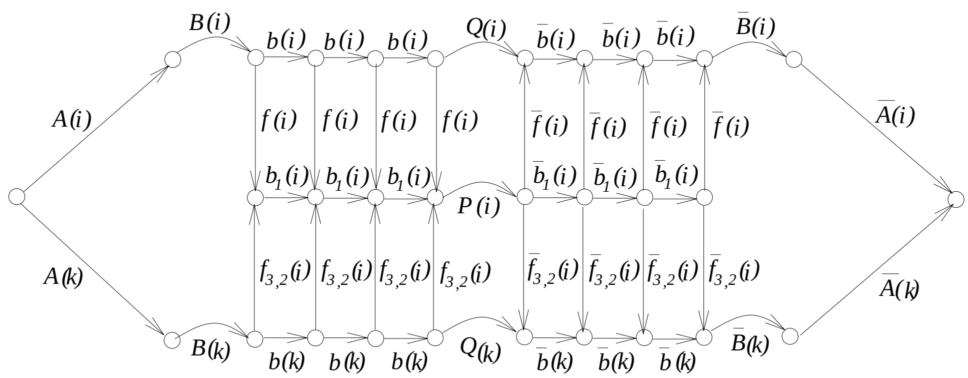

Fig. 6. $m=0,\ n=3$

Notice that in the first case there is a path in this automaton from the initial state to the terminal state labelled by $w(2,j,3)$ and in the second case there is such a path labelled by $w(0,k,3)$ (the lower boundaries of the graphs). Hence in each case $A(w(m,i,n))$ accepts the required word. Similar analysis applies in the other cases with corresponding sets of defining relations.

The basic idea involved in proving lemma 6.4 is to construct an inverse automaton $I(M)$ associated with the Minsky algorithm M and a configuration $(2^{m_0},1,0)$ where $m_0 \notin X$. This inverse automaton is such that the transition semigroup $T(M)$ of

$I(M)$ is a homomorphic image of the semigroup $S(M)$ constructed above and is in the variety $\hat{\mathcal{A}}$. In addition one shows that the word $w(2^{m_0}, 1, 0)$ labels a path which is not a loop in $I(M)$, and so it is not equal to 0 in $T(M)$. It follows immediately of course that $w(2^{m_0}, 1, 0) \neq 0$ in $S_{\mathcal{B}}(M)$ for every variety $\mathcal{B}$ containing $\hat{\mathcal{A}}$ and hence that Lemma 6.4 holds.

The idea is that the underlying graph of $I(M)$ should contain the Schützenberger automaton of the canonical word $w(2^{m_0}, 1, 0)$ and have the following properties.

a) The automaton $I(M)$ is closed under application of all of the defining relations of the semigroup $S(M)$. That is, if in $I(M)$ there are two vertices p and q and a path starting at p and ending at q labelled by one side of one of the relations then there is also a path in $I(M)$ from p to q labelled by the other side of that relation. From this it follows that the transition semigroup $T(M)$ of $I(M)$ is a homomorphic image of $S(M)$.

b) The transition semigroup $T(M)$ of $I(M)$ is in the variety $\hat{\mathcal{A}}$.

We refer to the paper [68] for full details.

We note that the semigroup $S(M)$ itself is not in the variety $\hat{\mathcal{A}}$. For example, the word $A(i)a(i)A(i)^{-1}a(i)^{-1}$ is not an idempotent in $S(M)$ since its Schützenberger automaton is linear (no relations apply), but the word is 1 in the free Abelian group. Hence by Theorem 6.2, $S(M) \notin \hat{\mathcal{A}}$. We can add the relations like

$$A(i)a(i)A(i)^{-1}a(i)^{-1} = (A(i)a(i)A(i)^{-1}a(i)^{-1})^2$$

to $S(M)$, but still there will be infinitely many (more sophisticated) commutators which will not be idempotents in the resulting semigroup. Thus we are unable to push $S(M)$ into $\hat{\mathcal{A}}$ by adding only finitely many relations. This raises the following question.

Problem 6.1 *Is there a finitely presented inverse semigroup with undecidable word problem, which belongs to* $\hat{\mathcal{A}}$*?*

Finally we raise the question as to whether or not the variety $\hat{\mathcal{A}}$ is a "bound" between decidability and undecidability of the word problem in the class of varieties of inverse semigroups.

Problem 6.2 *Is it true that every variety of inverse semigroups that is strictly contained in* $\hat{\mathcal{A}}$ *has decidable word problem? In particular, does the variety of inverse semigroups generated by the bicyclic semigroup have decidable word problem?*

In [66] we conjectured that the variety generated by the bicyclic semigroup has undecidable word problem. At this stage we are unable to either prove or disprove this conjecture. It is clear that this variety is properly contained in $\hat{\mathcal{A}}$.

References

1. S.I. Adian, *Defining relations and algorithmic problems for groups and semigroups*, Tr. Mat. Inst. Steklov 85 (Russian); Am. Math. Soc. (translation) 152, 1967.
2. S.I.Adian *Burnside problem and identities in groups*, Moscow, Nauka, 332 p. (1975), translated by John Lennox and J. Wiegold, Ergebnisse der Math and ihrer Grenzgebiete, 95, Springer-Verlag, 1979.

3. S.I.Adian, G.S.Makanin, *Studies in algorithmic questions of algebra*, Algebra, Mat. Logic, Number theory, Topology, Proc. Steklov Institute of Mathematics, Vol. 168, (1984), English translation (1986), issue 3 of 4, 207–226.
4. M.D. Atkinson, *Computational group theory,* Academic Press, 1984.
5. J. Avenhaus and K. Madlener, *The Nielsen reduction and P-complete problems in free groups*, Theor. Comp. Science, 32(1984), 61-76.
6. L.Babai, E.Luks and A.Seress, Permutation groups in NC, Proc. 19th ACM STOC (1987), 409-420
7. A.M.Ballantyne and D.S.Lankford, *New decision algorithms for finitely presented commutative semigroups,* Computers and Math. with Appl., 7 (1981), 159-165.
8. D.B.Bean, A.Ehrenfeucht, and G.McNulty, *Avoidable patterns in strings of symbols,* Pacific J. Math., 85 (1979), no. 2, 261-294.
9. C. Bennett *Time/Space tradeoffs for reversible computations*, SIAM J. of Computing, 18 (1989), 766-776.
10. J. Berstel *Transductions and Context-Free Languages*, Teubner Studienbucher, 1979.
11. J-C. Birget, S.Margolis, J.Meakin, *The word problem for inverse monoids presented by one idempotent relator*, Theoretical Comuter Science, 1993 (to appear).
12. J-C. Birget, S. W Margolis, J. C.Meakin and P. Weil *PSPACE - complete problems for subgroups of free groups and inverse automata,* preprint, 1993.
13. A.P.Biryukov, *Some algorithmic problems for finitely defined commutative semigroups,* Siberian Math. J., 8 (1967), 384-391.
14. L.A.Bokut' *The algorithmic problems and the embedding theorems: some open questions for rings, groups and semigroups*, Izv. Vyssh. Uchebn. Zaved. Mat., (1982), No. 11, 3–11.
15. L.A.Bokut', G. P. Kukin, *Undecidable algorithmic problems for semigroups, groups and rings*, Algebra, Topology, Geometry, Vol. 25. Itogi Nauki i Techniki, 3–66, (1987).
16. W.W.Boone, *Certain simple unsolvable problems in group theory I-VI*, Proc. Kon. ned. akad. Wetensch. A, 57 (1954) 231-237, 492-497, 58 (1955), 252-256, 571-577, 60 (1957), 22-27, 227-232.
17. B.Buchberger and R.Loos, *Algebraic simplification*, in "Computer Algebra. Symbolic and algebraic computation. B.Buchberger, G.E.Collins, and R.Loos (ed.)", Springer-Verlag, 1982, 23-65.
18. S. Burris, H.P. Sankappanavar *A Course in Universal Algebra*, Grad. Texts in Math 78, Spriner-Verlag, 1981.
19. I. M. Chiswell, D. J. Collins and J. Huebschmann *Aspherical group presentations,* Math. Z. 261 (1981), 1 - 36.
20. S. Cho, D. Huynh, *Finite-automaton aperiodicity is PSPACE-complete*, Theor. Comp. Sci. 88 (1991), 99-116.
21. J. R. Cho and S. J. Pride *Embedding semigroups into groups, and the asphericity of semigroups,* Int. J. Algebra and Computation, Vol. 3, No. 1, (1993), 1 - 15.
22. A.H.Clifford and G.B.Preston, *The algebraic theory of semigroups. v. 1, 2*, Amer. Math. Soc., Providence, 1961.
23. D. E. Cohen, *Combinatorial Group Theory: a Topological approach,* Lond. Math. Soc. Student Texts 14 (1989).
24. D.Cowan, *A class of varieties of inverse semigroups*, J. of Algebra 141(1991), 115-142.
25. M. Dehn, *Über unendliche diskontinuerliche Gruppen,* Math. Ann., 1912, Bd. 71, S. 116-144.
26. L.E.Dickson, *Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors,* Amer. J. Math. 35 (1913), 413-422.
27. S.Eilenberg, *Automata, languages and machines,* Pure and Applied Mathematics, Vol. B, Academic Press, 1976.
28. V.A.Emelichev, *Commutative semigroups with one defining relation,* Uchenye zap. Shuiskogo ped. in-ta, 6 (1958), 227-242.
29. D.B.A.Epstein , J.W. Cannon, D.F. Holt, S.V.F.Levy, M.S. Patterson and W.P. Thurston *Word Processing in Groups*, Jones and Bartlett, Boston, 1992.
30. J. Fountain *Abundant semigroups*, Proc. London Math. Soc. 44 (1982), 103-129.
31. H.Furstenberg, *Poincaré recurrence and number theory*, Bull. Amer. Math. Soc. 5 (1981), 211-234.
32. S. M. Gersten *Intersections of finitely generated subgroups of free groups and resolutions of graphs,* Inv. Math. 71 (1983), 567 - 591.

33. L. M. Gluskin *Elementary generalized groups,* Mat. Sbornik 41 (1957), 23 - 37 (in Russian).
34. M. Gromov, *Hyperbolic groups,* IHES, France, Preprint. P. 187.
35. Yu.Sh Gurevich, *The problem of equality of words for certain classes of semigroups*, Algebra i Logika, 25–35.
36. M. Hall Jr., *A topology for free groups and related groups,* Ann. of Math. 52 (1950), 127-139.
37. P.Hall, *On the finiteness of certain soluble groups*, Proc. London Math. Soc. Vol. 3, No. 9, 595–622 (1959).
38. G.Hermann, *Die Frage der enlich vielen Schritte in der Theorie der Polynomideale,* Math. Ann. 95 (1926), 736-788.
39. Peter Higgins *Techniques of semigroup theory,* Oxford Univ. Press (1992).
40. Phillip Higgins, *Categories and Groupoids,* Van Nostrand - Rheinhold, New York (1971).
41. Sergei Ivanov *Relation modules and relation bimodules of groups, semigroups and associative algebras,* Int. J. Algebra and Computation, Vol. 1, No. 1, (1991) 89 - 114.
42. Sergei Ivanov, *The Burnside problem for all sufficiently large exponents*, Int'l Journal of Algebra and Computation, to appear.
43. P. Jones, S. Margolis, J. Meakin, J. Stephen *Free products of inverse semigroups II*, Glasgow Math J. 33 (1991) 373 - 387.
44. O.G.Kharlampovich, *The word problem for subvarieties of the variety* $\mathcal{N}_2\mathcal{A}$, Alg. i Logika, Vol. 26, 481–501 (1987); translation in Algebra and Logic, Vol. 26 (1987).
45. O.Kharlampovich and M.Sapir, *Algorithmic problems in varieties of groups, semigroups, associative and Lie algebras*, in preparation.
46. O.G.Kharlampovich, M.Sapir, *Non-resudually finite relatively finitely presented groups in the variety* $\mathcal{N}_2\mathcal{A}$., preprint.
47. E.I.Kleiman, *On bases of identities of Brandt semigroups*, Semigroup Forum, 13, 3 (1977), 209-218.
48. D. Kozen *Lower bounds for natural proof systems*, Proc. 18th Annual Symosium on Foundations of Computing (FOCS) (1977), 254-266.
49. G.Lallement, *Semigroups and combinatorial applications,* Wiley & Sons, New York, 1979.
50. Roger C. Lyndon and Paul E. Schupp, *Combinatorial group theory,* Springer-Verlag, 1977.
51. I.G.Lysenok *The infinitude of Burnside groups of period* 2^k *for* $k \geq 13$, Russian Math. Surveys, 47 (1992), 2, 229-230.
52. A.I.Mal'cev, *On the homomorphisms onto finite groups*, Uchebn. Zapiski Ivanovsk. ped.instituta, 1958, V. 18, No. 5, 49–60; also in "Selected papers", Vol. 1, Algebra, (1976), 450–461.
53. A.I.Mal'cev, *Algorithms and recursive functions*, Nauka, Moscow (1965), 391 pp.
54. S.Margolis and J.Meakin, *E-unitary inverse monoids and the Cayley graph of a group presentation,* J.Pure Appl. Algebra 58 (1989), 45-76.
55. S.W.Margolis, J.C. Meakin *Inverse monoids, trees ans context-free languages*, Transactions of the AMS, V. 335, No.1, 1993, 259–276.
56. S. W. Margolis, J. C. Meakin *Free inverse monoids and graph immersions* Int. J Algebra and Computation, Vol. 3, No. 1, (1993) 79 - 99.
57. S. Margolis, J.Meakin, J.Stephen, *Some decision problems for inverse monoid presentations,* in "Semigroups and their applications", D. Reidel, 1987, 99-110.
58. S. Margolis, J.Meakin, J.Stephen, *Free objects in certain varieties of inverse semigroups,* Canad. J. Math. XLII, 6 (1990), 1084-1097.
59. S. MacLane *Categories for the working mathematician,* Grad. Texts Math. No. 5, Springer, New York (1971).
60. S.Margolis, M.V.Sapir, *Quasi-identities of finite semigroups and symbolic dynamics,* to appear.
61. A. Markov, On the impossibility of certain algorithms in the theory of associative systems, Dokl. Acad. Sci. URSS 55 (1947), 583–586.
62. Yu. V. Matiyasevich *Studies in certain algorithmic problems in algebra and number theory*, Algebra, Mat. Logic, Number theory, Topology, Proc. Steklov Institute of Mathematics, Vol. 168, (1984), English translation (1986), issue 3 of 4, 227–252.
63. J.McCammond, *The solution of the word problem for the relatively free semigroups satisfying* $x^a = x^{a+b}$ *with* $a \geq 6$, Int'l J. Algebra and Computation, 1 (1991), no. 1, 1-32.
64. J. C. C. McKinsey, The decision problem for some classes of sentences without quantifiers, J. Symbol. Log., Vol. 8, No. 3, 61–76 (1943).

65. G.McNulty, C.Shallon, *Inherently non-finitely based finite algebras,* Lecture Notes Math. 1004 (1982), 206-231.
66. J.Meakin, *An invitation to inverse semigroups,* Proc. Conf. on Ordered Structures and Algebra of Computer Languages, Hong Kong, June 1991, K.P Shum and P.C. Yuen (Ed), World Scientific 1993, 91 - 115.
67. J.Meakin, M. V. Sapir *Congruences on free monoids and submonoids of polycyclic monoids*, J. Austr. Math. Soc. (Series A), 54 (1993), 236-253.
68. J.Meakin, M. V. Sapir *On the word problem in varieties of inverse semigroups* to appear.
69. I.L.Mel'nichuk, *Solvability of problems of equality and divisibility in certain finitely based varieties of semigroups,* in "Properties of semigroups," Leningrad, 1984, 91-103.
70. I.L.Mel'nichuk, *Undecidability of problems of equality and divisibility in certain varieties of semigroups,* Algebra i Logika, 23 (1984), no. 4, 430-438.
71. I.L.Mel'nichuk, M. V. Sapir and O.G.Kharlampovich, *The word problem in varieties of semigroups, rings and Lie algebras*, Sibirsk. Mat. Zh. 1986, 27, 6, pp. 144-156 (translated from Russian into English by Plenum Corp. in Sib. Math. J.).
72. M.L.Minsky *Recursive unsolvability of Post's problem of "tag" and other topics in the theory of Turing Machnes* , Ann. of Math. (2) , Vol.74 (1961), 437–455.
73. M.L.Minsky *Computation: finite and infinite machines*, Prentice-Hall, Englewood Cliffs, N.J., 1967.
74. M. Morse, G. Hedlund, *Unending chess, symbolic dynamics and a problem in semigroups*, Duke Math. J., 11, (1944), 1–7.
75. W.D. Munn, *Free inverse semigroups,* Proc. Lond. Math. Soc. 30, 1974, 385-404.
76. V. L. Murskii, *Some examples of varieties of semigroups*, Mat. Zametki, 3 (1968), 663–670.
77. H. Neumann *Varieties of groups*, Ergebnisse der mathematik und ihrer grengebiete, Band 37, New Series, Vol.37 (1967), 192 pp.
78. P. S. Novikov, *On the algorithmical undecidability of the problem of identity,* Doklady AN SSSR, 4, 85 (1952), 709-719.
79. P.S. Novikov, S.I. Adian, *On infinite periodic groups, I, II, III,* Izv. Akad. Nauk SSSR, Ser. Mat., 1, 32 (1968), 212-244, 2, 32 (1968), 251-524, 3, 32 (1968), 709-731.
80. L.O'Carroll, *Embedding theorems for proper inverse semigroups*, J.of Algebra 42(1976), 26-40.
81. A.Yu.Ol'shanskii, *Conditional identities of finite groups*, Sib. Matem. J., 15 (1974), 6, 1409-1413 (Russian, English translation see in Sib. Matem. J., 1975, 1000-1003).
82. A.Yu. Ol'shanski, *Geometry of defining relations in groups,* Moscow, "Nauka", 1989 , 448 pp. (Russian).
83. A.Yu.Ol'shanskii *Hyperbolicity of groups with subquadratic isoperimetric inequality,* Int. J. Algebra and Computation, Vol. 1, No.3, (1991), 281 - 290.
84. F.Pastijn, *A representation of a semigroup by a semigroup of matrices over a group with zero*, Semigroup Forum, 10 (1975), 238-249.
85. F.Pastijn, *Inverse semigroup varieties generated by E-unitary inverse semigroups,* Semigroup Forum, 24 (1982), 87-88.
86. Peter Perkins, *Decision problems for equational theories of semigroups*, Ph.D. Thesis, Univ. of California, Berkeley, 1966.
87. P.Perkins, *Basic questions for general algebras,* Algebra Universalis, 19 (1984), 16-23.
88. M. Petrich, *Inverse semigroups*, Wiley, New York, 1984.
89. M.Petrich, N.Reilly, *E-unitary covers and varieties of inverse semigroups,* Acta Sci. Math. Szeged, 46 (1983), 59-72.
90. J-E. Pin, *Varieties of formal languages,* Plenum, 1986.
91. E.L.Post, *Recursive unsolvability of a problem of Thue,* J. Symbolic Logic, 1, 12 (1947), 1-11.
92. G. B. Preston *Inverse semigroups,* J. London Math. Soc. 29 (1954), 396 - 403.
93. S.Pride *Low-Dimensional homotopy theory for monoids*, preprint, 1993.
94. N. Reilly, *Free combinatorial strict inverse semigroups*, J.London Math.Soc. 39 (1989), 102 –120.
95. V. N. Remeslennikov, V. A. Roman'kov *Model-theoretical and algorithmic questions in group theory*, Algebra, Topology, Geometry, Vol. 21, Akad. Nauk SSSR, Vsesoyuzn. Itogi Nauki i Techniki 3–79.
96. J. H. Remmers *On the geometry of semigroup presentations,* Adv. Math. 36 (1980), 283 -

296.
97. N. S. Romanovskii *The embedding problem for abelian-by-nilpotent groups*, Siberian Math. Journal, Vol. 21, No.2, 170–174 (1980).
98. B.Rosenblat, *Diophantine theories of free inverse semigroups,* Transl. Siberian Math. J. 26 (1986), 101-107 (translation in pp. 860-865).
99. R. Ruyle, Ph.D. Thesis, University of Nebraska-Lincoln, to appear, 1994.
100. M.V.Sapir *On quasivarieties generated by finite semigroups,* Semigroup Forum, 20 (1980), 1, 73-88
101. M.V.Sapir, *Algorithmic problems in semigroup varieties*, Izv. Vyssh. Uchebn. Zaved. Mat. Matematica, 1985, 12, p. 71-74 (Translated from Russian into English by Allerton press).
102. M.V.Sapir, *Problems of Burnside type and the finite basis property in varieties of semigroups* , Izv. Akad. Nauk SSSR, ser. Mat., 1987, 51, 2, p. 319-340, translation in Math USSR-Izv, Vol. 30 (1988), No 2, 295–314.
103. M. V. Sapir *Inherently nonfinitely based finite semigroups*, Mat. Sb., 1987, 133, 2, p. 154-166 (translated from Russian into English by American MS).
104. M. V. Sapir *Sur la propriete de base finie pour les pseudovarietes de semigroupes finis*, C.R. Acad. Sci, Paris, 1988, 306, Serie I, pp. 705-708 (in English and French).
105. M.V.Sapir, Algorithmic problems in varieties of semigroups , Algebra i Logika, (1988), 27,4, p. 440-463; translation in Algebra and Logic, 27 (1988), No 4, 275–292.
106. M.V.Sapir, On the word problem in periodic group varieties, Int'l J. of Algebra Comput., Vol. 1, No. 1, 115–126 (1991).
107. M.V.Sapir *Finite semigroups in varieties*, Izv. Akad. Nauk SSSR, ser. Mat., 1991, 55, 5, p.245-260.
108. M. V. Sapir *Weak word problem for finite semigroups*, in "Monoids and semigroups with applications. Proceedings of the Berkeley Workshop in Monoids. Berkeley, 1989. J.Rhodes (ed.)", World Scientific, 1991, 206-219.
109. M.V.Sapir, *Identities of finite inverse semigroups,* Int'l Journal of Algebra and Computation, 3 (1993), no. 1, 115-124.
110. H. E. Scheiblich *Free inverse semigroups,* Proc. Amer. Math. Soc. 38 (1973) 1 - 7.
111. B.M. Schein, *Representations of generalized groups,* IZV., Vyss. Uchebn. Zav. Matem., issue 3 (1962), 164-176 (Russian).
112. B.M.Schein *Free inverse semigroups are not finitely presentable,* Acta. Math. Acad. Sci. Hung. 26 (1975), 41 - 52.
113. J.P. Serre, *Trees*, Springer-Verlag, 1980.
114. P.V. Silva *Rational languages and inverse monoid presentations*, Int. J. Algebra and Computation, Vol. 2, No. 2, 1992, 187 - 208.
115. L.N.Shevrin, M.V.Volkov, *Identities of semigroups*, Izv. vuzov. Matematika, 11 (1985), 3-47.
116. J. Stallings, *Topology of finite graphs,* Invent. Math. 71, 1983, 551-565.
117. J.B. Stephen, *Application of automata theory to presentations of monoids and inverse monoids,* Ph. D. Thesis, Univ. of Nebraska, 1987.
118. J.B. Stephen, *Presentations of inverse monoids,* J. Pure and Applied Algebra 63, 1990, 81-112.
119. M.A.Taiclin, *Algorithmic problems for commutative semigroups,* Soviet Math. Dokl. 9 (1968), 201-204.
120. H.Tietze, *Uber die topologischen Invarianten mehrdimensionaler Mannigfaltigkeiten* Monatsh. Math. Phys., 19 (1908), 1-118.
121. B.Tilson, *Categories as algebras: an essential ingredient in the theory of finite monoids,* J. Pure Appl. Algebra, 48 (1987), 83-198.
122. Axel Thue, *Über die gegenseitigen Lage gleicher Terlegewisser Zeichenreilren,* Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiania, 7 (1906), 1-22.
123. Axel Thue *Probleme über Veränderlungen von Zeichenreihen nach gegebenen Regeln, Kristiania Videnskapsselskapets Skr.* I: Mat.-Naturvid. Kl. 1914, no. 10; reprinted in his Selected mathematical papers, Universitetsforlaget, Oslo, 493–524 (1977).
124. V. V. Wagner *Generalized groups,* Dokl. Akad. Nauk. SSSR 84 (1952), 1119 - 1122 (in Russian).
125. A.I.Zimin, *Blocking sets of terms*, Matem. Sbornik, 119, 3 (1982), 363-375.