BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 4, Number 3, May 1981

CONTEXT-FREE LANGUAGES, GROUPS, THE THEORY OF ENDS, SECOND-ORDER LOGIC, TILING PROBLEMS, CELLULAR AUTOMATA, AND VECTOR ADDITION SYSTEMS

BY DAVID E. MULLER AND PAUL E. SCHUPP

It turns out that there exists a surprising connection between certain ideas from the areas listed in the title. In this announcement we try to briefly outline the connection and to state our principal results. Very roughly, we use concepts from formal language theory, group theory, and the theory of ends to investigate a class of graphs which we call context-free graphs. Using the results obtained and Rabin's theorem on the decidability of the monadic second-order theory of the infinite binary tree, we show that the monadic theory of any context-free graph is decidable. There are several classes of extensively investigated decision problems which are essentially problems on the grid of integer lattice points in n dimensions. We here have in mind various questions concerning tiling problems, cellular automata, and vector addition systems. Most of these problems are known to be unsolvable in the classical case. We show that these problems all make sense on a very general class of graphs and are all solvable on any context-free graph.

A finitely generated group can be described by a presentation $G = \langle X; R \rangle$ in terms of generators and defining relators. (All groups and presentations which we mention are assumed to be finitely generated.) The *word problem* of G is the set $W(G)$ of all words on the generators and their inverses which represent the identity element of G. Anisimov [1] raised the question: "If $W(G)$ is a context-free language in the sense of formal language theory, what can one say about the algebraic structure of G?" We were led to conjecture that G has context-free word problem if and only if G has a free subgroup of finite index, and we have essentially proven the conjecture.

Our main tool is the theory of ends. Let Γ be the Cayley graph of a finitely-generated group $G = \langle X; R \rangle$. Let $\Gamma^{(n)}$ denote the subgraph of Γ consisting of all vertices and edges connected to the identity by a path of length less than n. The *number of ends* of G is the limit as n goes to infinity of the number of infinite components of $\Gamma \backslash \Gamma^{(n)}$. Stallings [4] proved that a group with more than one end has a particular structure in terms of certain group-theoretic constructions. We prove that an infinite group with context-free word problem has

Received by the editors November 6, 1980.

1980 *Mathematics Subject Classification*. Primary 03D05, 20F10, 68F05; Secondary 03B25, 68D05.

more than one end. A group G is *accessible* if it cannot be decomposed arbitrarily often using the constructions mentioned in Stalling's theorem.

THEOREM 1. *A finitely generated group G is free if and only if G has context-free word problem and is torsion-free. A finitely generated group G has a free subgroup of finite index if and only if G has context-free word problem and is accessible.*

A *finitely-generated graph* is a connected labelled graph Γ with a vertex v_0 chosen as origin and with a fixed upper bound on the degrees of vertices. We assume that the labels on the edges of Γ come from a finite label alphabet Σ. As before, we use $\Gamma^{(n)}$ to denote the subgraph of Γ consisting of all vertices and edges connected to v_0 by a path of length less than n. If C is a connected component of $\Gamma\backslash\Gamma^{(n)}$ a *frontier point* of C is a vertex of C having distance n from v_0. If C is a component of $\Gamma\backslash\Gamma^{(n)}$ and C' is a component of $\Gamma\backslash\Gamma^{(m)}$ we say that C and C' are *end-isomorphic* if there is a label-preserving graph isomorphism between C and C' which takes frontier points to frontier points. A finitely-generated graph Γ is *context-free* if there are only finitely many isomorphism types under end-isomorphism. (For example, consider the infinite k-ary tree which is a tree with origin v_0 and k distinct edges labelled $\sigma_1, \ldots, \sigma_k$ leaving each vertex. There is only one isomorphism type since, regardless of n, a component of $\Gamma\backslash\Gamma^{(n)}$ is isomorphic to the whole tree.) The class of machines associated with the class of context-free languages is the class of pushdown automata. If M is a pushdown automaton it is possible to associate with M a finitely generated graph $\Gamma(M)$ whose vertex set is the set of possible total states of M and whose edges indicate transitions.

THEOREM II. *A graph Γ is context-free if and only if Γ is the graph of a pushdown automaton.*

A powerful positive result concerning decision problems in logic is Rabin's theorem [3] that the monadic second-order theory of the labelled n-ary tree is decidable. In this language individual variables represent vertices. If $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$ and w is a word on Σ and x is an individual variable, then xw represents the unique vertex obtained by starting at x and following the path with label w. One has set variables representing sets of vertices and the $\in$-relation of set membership. One forms sentences by using quantifiers and logical connectives in the usual way. The great power of the language is that one can quantify over sets.

One can define a very similar language for any finitely-generated graph Γ. If Γ is labelled from the finite alphabet Σ and w is a word on Σ, then xw denotes the set of vertices obtainable by starting at x and tracing out a path with label w.

Thus xw now denotes a finite set. We think of individual variables as denoting singleton sets and take the set inclusion relation, $\subseteq$, as basic.

THEOREM III. *If* Γ *is any context-free graph then the monadic second-order theory of* Γ *is decidable.*

Very roughly, one proves Theorem III in the following way. Questions about the monadic second-order theory of Γ are equivalent to questions about certain tiling problems on the graph Γ. Using heavily the fact that Γ is the graph of a pushdown automaton M of special type, one can reduce questions about tilings of Γ to questions about tilings of a tree T_Γ associated with M and then reduce the problems on T_Γ to problems about tiling a full k-ary tree S_Γ and apply Rabin's theorem to solve the problems.

In von Neumann's conception of a cellular automaton, there are identical finite state automata at the integer lattice points of n-dimensional space. A neighborhood set N is specified. (Without loss of generality, one can take N to consist of a vertex together with the other vertices connected to it by a single edge.) The next state of any particular automaton depends on the states of the automaton and its neighbors. Each automaton thus changes state according to the same *local transition function* δ. A *state of the universe* is an assignment of a state from Q to each automaton. Let D denote the set of all possible states of the universe. The local transition function δ induces a global transition function $\Delta: D \longrightarrow D$. One is interested in whether or not Δ is surjective or injective. Whether or not these questions are algorithmically decidable in the classic case has not been settled, but they are almost surely undecidable.

One can imagine a cellular automaton whose underlying universe is a context-free graph Γ. Suppose that N is a finite labelled graph with a vertex r distinguished as the *root* of N such that for each vertex v of Γ there is a unique labelled graph embedding $\phi_v: N \longrightarrow \Gamma$ with $\phi_v(r) = v$. Then we can consider cellular automata whose underlying universe is Γ and which have neighborhood set N.

THEOREM IV. *Let* Γ *be a context-free graph with neighborhood set* N. *Then there is an algorithm which decides, when given a state set and a local transition function, whether or not the global transition function is surjective* (*or injective*).

Finally, we turn to problems associated with vector addition systems. In the standard formulation, an n-dimensional vector addition system consists of a finite set U of n-dimensional integer vectors (i.e., integer n-tuples). Note that if $u \in U$, then $-u$ need not be in U. As usual, the first quadrant of $\mathbf{Z}^n$ consists of vectors all of whose entries are nonnegative. If z is a point of the first quadrant, the *reachability set* of z with respect to U, denoted by Uz, consists of the points v which can be obtained from z by successively adding vectors in U while staying within the first quadrant. Precisely, $v \in Uz$ if there exists a sequence $z = v_1, \ldots, v_k = v$ where each v_i is in the first quadrant, and, for $i = 1, \ldots, k-1$, there exists a $u_i \in U$ such that $v_{i+1} = v_i + u_i$. The *membership problem* for the vector addition system U asks for an algorithm which, when given points z and v of the first quadrant, decides if $v \in Uz$. It is not presently known if the membership problem is always solvable.

The *inclusion problem* for reachability sets asks for an algorithm which, when given two n-dimensional vector addition systems U_1 and U_2 and two points z_1 and z_2 in the first quadrant, decides whether or not $U_1 z_1 \subseteq U_2 z_2$. Rabin has proved that the inclusion problem is unsolvable. (See Hack [2].)

The idea of a vector addition system makes sense on an arbitrary finitely-generated graph Γ labelled by an alphabet Σ. Let t be a vertex of Γ and let W be a finite set of words of Σ. The *quadrant* tW of t with respect to W consists of all vertices t_n such that there is a finite sequence $v = t_1, t_2, \ldots, t_n$ where for each $i = 1, \ldots, n-1$, there exists a $w_i \in W$ with $t_{i+1} \in t_i w_i$ (where $t_i w_i$ denotes the set of all vertices obtainable by starting at t_i and tracing out a path with label w_i). A *vector addition system* on Γ consists of a vertex t and two finite sets W and U of words of Σ. If $z \in tW$, the reachability set Uz of z with respect to U, consists of all vertices s such that there is a sequence $z = s_1, \ldots, s_k = s$ with each $s_i \in tW$ and for each $i = 1, \ldots, k-1$ there is a $u_i \in U$ with $s_{i+1} \in s_i u_i$. The membership and inclusion problems can now be formulated in the obvious way.

THEOREM V. *Let Γ be a context-free graph. Then the membership and inclusion problems for vector addition systems on Γ are uniformly solvable.*

The proof of the last two theorems consists of showing that the problems are expressible as sentences in the monadic second-order theory of the underlying graph Γ.

REFERENCES

1. A. V. Anisimov, *Group languages*, Kibernetika **4** (1971), 18–24.

2. M. Hack, *The equality problem for vector addition systems is undecidable*, Theoret. Comput. Sci. **2** (1976), 77–96.

3. M. O. Rabin, *Decidability of second-order theories and automata on infinite trees*, Trans. Amer. Math. Soc. **141** (1969), 1–35.

4. J. Stallings, *Group theory and three-dimensional manifolds*, Yale Univ. Press, New Haven, Connecticut, 1971.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, ILLINOIS 61801