Franz-Josef Brandenburg , Jean Berstel,
Detlef Wotschke (editors):

Trends and Applications in Formal Language Theory

Dagstuhl-Seminar-Report; 23
14.10.-18.10.91 (9142)

ISSN 0940-1121

Tel.: +49-6871 - 2458
Fax: +49-6871 - 5942

Das Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI) ist eine gemeinnützige GmbH. Sie veranstaltet regelmäßig wissenschaftliche Seminare, welche nach Antrag der Tagungsleiter und Begutachtung durch das wissenschaftliche Direktorium mit persönlich eingeladenen Gästen durchgeführt werden.

Bezugsadresse: Geschäftsstelle Schloß Dagstuhl
Informatik, Bau 36
Universität des Saarlandes
W - 6600 Saarbrücken
Germany
Tel.: +49 -681 - 302 - 4396
Fax: +49 -681 - 302 - 4397
e-mail: office@dag.uni-sb.de

Jean Berstel
Franz-Josef Brandenburg
Detlef Wotschke
(editors)

Trends and Appplications in Formal Language Theory

Dagstuhl-Seminar
Oct. 14 - 18, 1991

Report
on the 1st Dagstuhl-seminar on
Trends and Applications in Formal Language Theory

Formal language theory is one of the classical and fundamental areas of Computer Science. For more than four decades a broad theoretical basis and a wide spectrum of applications of language theory has been developed, with the 1960´s and 1970´s as the "golden age" of formal language theory. Since then other areas became more attractive to many theoretical computer scientists and research on formal languages is often presented as part of other theories. Under this perspective the seminar on Trends and Applications in Formal Language Theory aims at serving as a forum for formal language theorists to make their research transparent as part of their theory.

The presented talks addressed a wide range of themes within formal language theory. Particular emphasis was laid upon the complexity of languages and their descriptions, i.e. both static and dynamic complexity theory. A more detailed classification can be done along the following keywords:

- description complexity and general complexity
- nondeterminism, unambiguity, determinism
- generation of complex objects
- special use of regular and context-free methods
- parallel rewriting and codes

Besides the 27 presented talks there were a lot of discussions is small goups and a general one focussing on perspectives of formal language theory. Among the participants there was a broad consensus that the visibility of formal language theory must be improved. Activities in this direction of some participants were gratefully acknowledged. Clearly, the best support would come from a regular conference on formal language theory or a devoted section within an established conference series, with Dagstuhl seminars as a first step.

All participants appreciated the stimulating Dagstuhl atmosphere and wished to thank the staff for doing a perfect job.

Franz J. Brandenburg

Participants

Jürgen Albert, Würzburg
Jean Berstel, Paris/Frankreich
Franz-Josef Brandenburg, Passau
Anne-Brüggemann-Klein, Freiburg
Didier Caucal, Rennes/Frankreich
Bruno Courcelle, Bordeaux/Frankreich
Carsten Damm, München
Jürgen Dassow, Magdeburg
Jürgen Duske, Hannover
Joaquim Gabarro, Barcelona/Spanien
Massimiliano Goldwurm, Mailand/Italien
Josef Gruska, Hamburg
Günter Hotz, Saarbrücken
Dung T. Huynh, Dallas/USA
Birgit Jenner, München
Helmut Jürgensen, London/Ontario/Kanada
Alica Kelemenova, Bratislava/CSFR
Werner Kuich, Wien/Österreich
Klaus-Jörn Lange, München
Hing Leung, Frankfurt
Jorma Rissanen, San José/USA
Jürgen Schäfer, Frankfurt
Helmut Seidl, Saarbrücken
Ray Solomonoff, Saarbrücken
Wolfgang Thomas, Kiel
Paul Vitanyi, Amsterdam/Niederlande
Hermann Walter, Darmstadt
Dietmar Wätjen, Braunschweig
Andreas Weber, Frankfurt
Detlef Wotschke, Frankfurt

Abstracts

Halbringautomaten und TOL-Systeme

Werner Kuich, Technische Universität Wien

Ausgangspunkt sind TOL-Systeme, deren Parallelableitungen durch Kontrollsprachen reguliert werden. Diese werden dadurch verallgemeinert, daß formale Potenzreihen über endlichen Substitutionen auf formale Potenzreihen über Wörtern angewendet werden. Stammen die Potenzreihen aus abstrakten Familien von Potenzreihen (AFP), so erhält man unter gewissen Bedingungen - als Verallgemeinerung des AFL-Resultates für ETOL-Sprachen - wieder eine AFP.

Beschränkt man sich auf algebraische Potenzreihen über $\mathbb{N}^\infty$ und auf Homomorphismen, kann man - in Verallgemeinerung von Paz, Salomaa (1973) - Wachstumsmatrizen definieren und damit algebraische Potenzreihen (im Sinne der Funktionentheorie) erzeugen, die für gewisse kombinatorische Abzählprobleme eingesetzt werden können.

Monadic second-order definable graph transductions

Bruno Courcelle, Université Bordeaux

Formulas of monadic second-order logic can be used to specify graph transductions, i.e., multivalued functions from graphs to graphs. We obtain in this way classes of graph transductions, called monadic second-order definable graph transductions (or more simply definable transductions) that are closed under composition and preserve the two classes of context-free sets of graphs, namely the class of Hyperedge Replacement (HR) and the class of Vertex Replacement (VR) sets of graphs. These two classes can be characterized in terms of definable transductions and recognizable sets of finite trees. This characterization is independent of any rewriting mechanism like the ones upon which the definitions of HR and VR grammars are based. When restricted to words, the definable transductions are strictly more powerful than the rational transductions with finite image; they do not preserve context-free languages. We also describe the sets of discrete (edgeless) labeled graphs that are the images of HR and VR

sets under definable transductions: this gives a version of Parikh's theorem (i.e., the characterization of the commutative images of context-free languages) which extends the classical one and applies to HR and VR sets of graphs.

Complexities and Formal Languages - Open Questions and some new Results
Klaus-Jörn Lange, TU München

The very deep connections between the theory of formal languages and complexity theory are illustrated by some typical relations between the two areas. The first part of the talk considers complexity measures defined by sequential models. Using Engelfriet's model of Automata with Abstract Storage certain relations concerning membership and emptiness problems are derivable in a very general way. The investigation of the closure properties of complexity classes leads to complexities of formal language operations, which are relative to pairs of complexity classes and characterize some long standing open problems. The second part deals with parallel complexities. After introducing some basic models and their relations to sequential classes we give more precise localizations of formal language problems within the framework of complexity classes. On the other hand, it is possible to get new characterizations of parallel models like CREW and EREW PRAMs by the application of some methods and notions well-known from formal language theory. Finally some new representations of the classes NC^k and DSPACE(log n) in terms of CRCW PRAMs are obtained using notions and criteria, which were developped in order to classify those PRAM algorithms which are effeciently implementable on existing parallel machines.

Validating SGML document grammars
Anne Brüggemann-Klein, Universität Freiburg
(joint work with Derick Wood, University of Waterloo)

A document grammar in the sense of the ISO standard 8879 on the Standard Generalized Markup Language (SGML) is essentially a context-free grammar with rules of the form $A \rightarrow [_A\ E\ ]_A$ with unique brackets $[_A$ and $]_A$ and a regular expression E. The expression E must be deterministic, i.e. each word can be "matched" by E in a unique way with one symbol of look-ahead.

In this talk we investigate the class of deterministic regular languages, i.e. of the languages that can be denoted by deterministic regular expressions.
We show the following Theorem: Given a DFA M, we can decide in time quadratic in the size of M whether the language of M is a deterministic regular language.
If the language of M is deterministic, a deterministic expression denoting it can be constructed from M.

Different Limitations of Lindenmayer Systems

Dietmar Wätjen, TU Braunschweig

We consider limited ETOL systems where at each step of the derivation process exactly
(a) min $\{k, \#_a w\}$ occurrences ($\#_a w$ being the number of occurrences of a in w)
(b) min $\{k, |w|\}$ symbols
of the word considered have to be rewritten ((a) k-limited or (b) uniformly k-limited ETOL systems). If the value of the limitation is given by a function $f: \mathbb{N} \rightarrow \mathbb{N}_0$ where, beginning from the axiom, f(1) is the limitation in the first step, f(2) the limitation in the second step etc., then we get the definitions of function-f-limited or uniformly function-f-limited ETOL systems.
In the non-extended case (limited TOL systems) we get a lot of incomparability results. Especially, for different values of k, the families of (uniformly) k-limited TOL languages are incomparable to each other.
In the extended case we get, among others, that the family of ETOL languages is strictly included in the family of k-limited ETOL languages. It can be shown that every k-limited ETOL language can be generated by a k-limited ETOL system with 2 tables, but also by a system with at most 3 active symbols in each table.

On the contour word of polyominoes

Danièle Beauquier, Université Paris (delivered by Jean Berstel)

A polyomino is a finite union of unit cells that is homeomorphic to a disk. It can also be defined by its contour word; the set of contour words, denoted by SC, is the set of words $u \in C$, that have no proper factor in C, where C is the set of words w such that

$$|w|_x = |w|_{\bar{x}} \text{ and } |w|_y = |w|_{\bar{y}}.$$

The following are shown: Given a regular language R,

1) $R \cap C \neq \emptyset$ is decidable (due to M. Latteux)

2) $R \cap SC \neq \emptyset$ is undecidable (D. Beauquier)

3) $R \cap CSC \neq \emptyset$ is decidable (D. Beauquier, M. Latteux, and K. Slowinski)

Here CSC denotes the set of contour words of convex polyominoes.

A note on the degree of nondeterminism

Jürgen Dassow, TU Magdeburg (joint work with Henning Bordihn, TU Magdeburg)

Let $G = (V, P, w)$ be a TOL system. Then the degree of nondeterminism $\det(G)$ introduced by G. Rozenberg is defined as $\det(G) = \max_{p \in P} \max_{x \in V} \#\{z_x : x \to z_x \in P\}$.

For a TOL language L, we set $\det(L) = \min\{\det(G) : L = L(G)\}$.

Rozenberg has shown that, for any integer n, there is a TOL language L_n, such that $\det(L_n) = n$ and that $\det(L) \leq 2$ holds for any ETOL language L. Clearly, an (E)TOL language is an (E)DTOL language if and only if $\det(L) = 1$.

This measure has also been studied for k-limited L systems by Wätjen and Dassow, and it was shown that $\det(L) \leq k$ for all k-limited L languages and that, for any $r \leq k$, there is a k-limited language L_r such that $\det(L_r) = r$.

First we mention a corresponding result for pure context-free grammars (or equivalently, OS systems).

Theorem. i) For any $n \geq 1$, there is a pure context-free language such that $\det(L_n) = n$.

ii) For any context-free language, $\det(L) \leq 2$.
Moreover, $\det(L) = 1$ holds if and only if L consists of a single word or L is empty.

Analogous statements hold for other types of pure languages.

Further we discuss decision problems concerning the degree of nondeterminism.

Theorem. For given TOL systems (k-limited L systems and pure context-free grammars) G and given integer $n \geq 1$, it is undecidable whether or not $\det(L(G)) = n$ and whether or not $\det(L(G)) = \det(G)$?

There is no algorithm for the computation of $\det(L(G))$ and of a TOL system (k-limited L system, pure context-free grammar) G' with $L(G') = L(G)$ and $\det(G') = \det(L(G))$.

The parallel complexity of some problems on formal languages and automata theory

Joaquim Gabarró, LSI - UPC, Barcelona

We analyse the complexity of some problems from formal languages and automata theory. We take as model of computation the PRAM and the NC complexity class.

-First we study the word problem over free partially commutative groups introduced by C. Wrathall. Our approach uses dependency graphs and associates to every group a rewriting system over these graphs. We apply there ideas to find an NC algorithm.

Second we consider the word problem for free partially commutative groups. To study this problem we use first order logic and majority quantifiers. It is proved that it belongs to an uniform class TC^*.

Third we take the membership problem for Petri nets. As above this problem belongs to TC^* class.

Finally we see, that the complexity of computing bisimulations is sequential. To do this we reduce the circuit value problem for monotonic alternating circuits to strong-bisimilarity.

On the complexity of some finite-state automaton problems

Birgit Jenner, TU München (joint work with Carme Àlvarez, UPC Barcelona)

We consider the complexity of the following functional variants of the non-emptiness problem for nondeterministic (NFA) or deterministic (DFA) finite-state automata:

(i) the ranking problem for NFA:
Given an encoding of an NFA M and a word $x \in \Sigma^*$, compute the number of words $y \in L(M)$ with $y \leq x$;

(ii) the ranking problem for DFA;

(iii) the maximal word problem für NFA:
Given an encoding of an NFA M and a word $x \in \Sigma^*$, compute the maximal word $y \in L(M)$ with $y \leq x$.

(iv) the maximal word problem for DFA.

We show, that the last problem is NSPACE(log n)-complete, while the first three are, respectively, complete for the counting and optimization classes spanL, #L, and optL, that we define analogously to their polynomial-time counterparts spanP, #P, and optP: Functions

in #L count the number of accepting computations of an NSPACE(log n)-machine, functions in spanL count the number of different output values of such a machine with additional output tape, and functions in optL compute the maximum of all output values. We place the three classes in the setting of other complexity classes and show that spanL and #P are very similar: they share all their complete functions with respect to log-space metric-reducibility. Consequently, the ranking problem for NFA is computable in polynomial time iff #P-functions and all problems in the polynomial hierarchy are computable in polynomial time. This reveals the hardness of spanL-functions. On the other hand, #L and optL-functions, and in particular the ranking problem for DFA and the optimal word problem for NFA, are easy. They are not only computable in polynomial time, but even "efficiently parallelizable": $\#L \subseteq NC^2$ and optL $\subseteq AC^1$. Matrix powering of positive integer matrices turns out to be #L-complete and (MAX, concat) word matrix powering optL-complete.

Buffer-Normal-Forms and synchronization of DPDAs

Jürgen Albert, Universität Würzburg
(joint work with Karel Culik II, Columbia and Juhani Karhumäki, Turku)

We define simplifications of deterministic pushdown automata as a family of bijective DGSM-mappings on the set of accepting computations. These mappings are used to remove inessential stack-actions and to achieve several Normal-Form-properties of DPDA, e.g.

- faithfulness
- no pushing ε-moves
- no finite pushing
- reduced DPDAs.

It is conjectured, that a confluence property holds for simplifications and subcases are given, where this confluence relation is fulfilled.

Based on this claim, a second Conjecture can be concluded, namely: For synchronizable and equivalent DPDAs M_1, M_2 there exists a common simplification M' for both M_1 and M_2.

Descriptions and descriptional complexity

Josef Gruska, Universität Hamburg and Slovak Academy of Sciences, Bratislava

Informatics is usually viewed as dealing primarily with computational and information processing problems, in general. Computational complexity (broadly unterstood) is usually seen as perhaps the most mature part of Informatics with significant impacts on mathematics. An attempt has been made in the talk to demonstrate that actually and to an increasingly large extent Informatics is dealing with problems related to descriptions (representations, specifications, declarations, languages, etc.). Descriptional complexity (broadly understood) should then be seen as an important part of complexity theory, closely related and complementary to computational complexity, with important impacts on foundations of other sciences.

Decomposing a k-valued transducer into k unambiguous ones

Andreas Weber, Universität Frankfurt

We investigate the inner structure of finite-valued finite transducers. We show: A k-valued nondeterministic generalized sequential machine (NGSM) M can be effectively decomposed into k unambiguous NGSMs $M_1,\ldots, M_k$ such that the transduction realized by M is the union of the transductions realized by $M_1,\ldots, M_k$. Each M_i is of double exponential size and can be constructed in deterministic double exponential time. By reduction, this result can be extended to normalized finite transducers (NFTs). As a consequence, the k-valued and the k-ambiguous NGSMs (NFTs) have the same generative power.

Influence of grammar classes to an effective description of languages

Alica Kelemenová, Slovak Academy of Sciences, Bratislava

The following results are presented:

1) For a language L, the total length of the optimal representation by CF grammars is bounded by the square of the length of its optimal representation by adult OL-systems.

2) For an infinite grammar form G and the corresponding class of grammars $\mathfrak{G}_x(G)$ and languages $\mathfrak{L}_x(G)$, where x is either strict or a general interpretation, we show that for an

arbitrary natural number n there is a language in $\mathcal{L}_X(G)$ with complexity at least n (with respect to $\mathcal{G}_X(G)$).

3) A colony of grammars is a simple grammar system consisting of regular grammars producing finite languages only and having the generative power of c.f. languages. For an arbitrary natural number k, there is a language which is determined by a k-element colony but by no (k-1) element colony.

Graph acceptors and first-order logic

Wolfgang Thomas, Universität Kiel

We introduce a notion of finite state graph acceptor, which can specify "projections of local properties" and whose expressive power matches a natural fragment of monadic second-order logic.

The considered graphs have labelled vertices and edges (with label alphabets A, B, resp.) and are of uniformly bounded degree $\leq k$. A graph acceptor is given by a finite state set Q, a finite set Δ of transitions (finite connected graphs with vertex labels in $A \times Q$, edge labels in B, and a designated center vertex), and a "constraint" (a boolean combination of conditions "there are $\leq n$ occurrences of transition τ"). A graph G with vertex set V is accepted if there exists a "tiling t by transitions", assigning to each vertex v a transition centered at v, such that for some $m \geq 0$: (1) the "m-sphere" around any $v \in V$ matches the transition of t centered at v, (2) the Q-components of the transitions of t define a consistent run $\rho: V \to Q$, (3) constraint C is satisfied for t.

Remark: Restricted to trees or words, these graph acceptors are equivalent in expressive power to finite (tree-)automata. (Here the constraints are superfluous.)

Theorem:

For any set L of finite graphs (of bounded degree $\leq k$):

(a) L is recognizable by a graph acceptor iff L is definable in existential monadic second-order logic.

(b) L is recognizable by a single-state graph acceptor iff L is definable in first-order logic.

In the proof we use the Ehrenfeucht-Fraissé-game technique of first-order model theory, in a form due to Hanf (1965).

The complexity of Deciding Readiness & Failure Equivalence for Processes

Dung T. Huynh, University of Texas, Dallas (joint work with Lu Tian)

In this talk we present some complexity results concerning the problem of deciding readiness and failure equivalences for finite state processes, and recursively defined processes specified by normed context-free grammars (CFGs) in Greibach normal form (GNF). The results are as follows. Readiness and failure equivalences for processes specified by normed GNF CFGs are both undecidable. For this class of processes, the regularity problem with respect to failure and readiness equivalence is also undecidable. Moreover, all these undecidability results hold even for locally unary processes. In the unary case, these problems are Π_2^P - complete. We also show that with respect to bisimulation equivalence, the regularity problem for normed GNF CFG processes is NL-complete. Readiness and failure equivalences for finite state processes are PSPACE-complete. This holds even for locally unary finite state processes. These equivalences are co-NP-complete for unary finite state processes. Further, for acyclic finite state processes, readiness and failure equivalences are co-NP-complete and they are NL-complete in the unary case. For finite tree processes, all equivalences are L-complete.

On a characterization of finite automata with limited nondeterminism

Hing Leung, Universität Frankfurt and New Mexico State University

In this study we want to design a new algorithmic solution to the limitedness problem on the amount of nondeterminism in NFA. This problem is a special case of the limitedness problem on distance automata, which has been solved independently by Hashiguchi, Leung and Simon. However, the algorithm to the general problem runs in exponential time and offers a double exponential upper bound on the distance when the automaton is limited in distance.

Using a new technique, we are able to derive a new characterization of finite automata with limited nondeterminism. From this, we deduce that if the amount of nondeterminism is bounded then it must be $\leq 2^n - 2$ with the assumption that only binary branching is allowed. We conjecture that the bound could be improved to O(n).

We believe that such a new technique may be extended to the study of limitedness problem for distance automata and provide, hopefully, a single exponential upper bound on distance which would then imply that the limitedness problem is in PSPACE.

Rewriting systems and monadic theory

Didier Caucal, IRISA, Rennes

We consider the monadic second-order theory of prefix replacement according to word rewriting systems and term context-free grammars.
The graph of the accessible prefix transitions from a given axiom and according to a labelled word rewriting system is effectively the transition graph of a pushdown automaton, and the converse is true. The monadic theory of the transition graph of a pushdown automaton is decidable (Muller and Schupp). Then the monadic theory is decidable for the prefix replacement of a word rewriting system, starting from an axiom. If we do not take into account the existence of an axiom, the monadic theory remains decidable because the prefix transition graph (not accessible) of a word rewriting system is effectively an equational graph (in the sense of Courcelle).
The monadic theory is undecidable for the prefix replacement according to linear context-free grammars on terms, and to ground term rewriting systems. In spite of these strongly negative results, we introduce a new notion on terms with variable. A term t is entire if any subterm of t is either a variable, or without variable, or has the same variables as t. We show that the class of accessible prefix transition graphs of labelled entire term of a grammar is exactly and effectively the class of rooted equational graphs of Finite out.degree (the in.degree may be Finite). Then, the monadic theory is decidable for the prefix replacement according to an entire term of a grammar, starting with an axiom. This decidability result must be extended to the non existence of an axiom. But and on the contrary to the word case, the prefix transition graph (not accessible) of an entire of a grammar is generally not equational.

A Natural Generalization of Context-free Languages

Günter Hotz, Universität Saarbrücken (joint work with Yongquang Guan)

We define c.f. net languages in the framework of free x-categories. These languages are special c.f. graph languages (planar). They are closed under concatenation, star and union. It holds Parikh's theorem and an analog to the group invariant of the author. In the case of these languages it is not a group but a groupoid.
The natural generalization of the c.f. word languages we get by multilinear interpretations of the x-category applying it to the c.f. net languages. These languages have the same closure

properties as the net languages and the Parikh theorem holds, too. These languages form an infinite hierarchy depending from the maximal degree (maximal sum of inputs and outputs of generaters of the x-category).
For k=1 one has the tree adjoint grammars (TAG's) of Joski. We give a simple procedure to decide the word problem for the case k=1 in time $O(n^6)$. We conjecture $O(n^{4+2k})$ in the general case. A more sophisticated algorithm of Guan decides the word problem in time $O(n^5)$ in case k=1.

Application of Formal Languages to Machine-Learning

Ray Solomonoff, Universität Saarbrücken

Prediction, accuracy of prediction, sample size of data needed, computational cost, and incremental learning are the basis of criticism of other work in Machine Learning.
A simple probablistic CFG is used to generate candidate solutions for the problem of inferring algebraic notation from numerical examples. The technique is extended to learning to solve linear algebraic equations and inferring the three laws of algebra.
Koza's Genetic Algorithm for general problem solving is described. We show how algorithmic probability can be used to improve his technique for generating new candidate solutions.

A New Approach to Formal Language Theory Using Kolmogorov Complexity

Paul Vitányi, CWI and University of Amsterdam (joint work with Ming Li)

It is feasible to reconstruct parts of formal language theory using algorithmic complexity (Kolmogorov complexity). We prove theorems on how to use Kolmogorov complexity, as a concrete, powerful, tool. We do not just want to introduce fancy mathematics; our goal is to help our readers (or listeners as the case may be) to do proofs in formal language theory in the most essential, usually easiest, sometimes even obvious ways. We give an alternative to pumping Lemma's and a new characterization of regular languages in terms of Kolmogorov complexity. We give a new quite general method to separate deterministic context free languages and nondeterministic context free languages. We illustrate the use of the new techniques through many examples which were considered difficult before. The approach is

also successful at the high end of the Chomsky hierarchy since one can quantify nonrecursiveness in terms of Kolmogorov complexity.

Hierarchies of Codes

Helmut Jürgensen, University of Western Ontario (joint work with S.S. Yu)

It has been observed, that many natural classes of codes can be characterized as sets of incompatible elements with respect to a binary relation on the free monoid. Indeed, very often this relation is actually a partial ordering and, moreover, expresses certain information theoretic properties concerning error resistance, decodability, synchronizability, etc. It turns and that the proper setting for this correspondence is abstract dependence theory in the sense of Cohn ("Universal Algebra"; see also Gécseg and Jürgensen, in Algebra Universalis, to appear). A "gap theorem" explains the step within the hierarchy of codes.
Another theorem characterizes when certain subhierarchies collapse.

Tree Automata with Cost Functions, II: Polynomial Boundedness

Helmut Seidl, Universität Saarbrücken

Cost functions for finite tree automata are mappings from transitions to (tuples of) polynomials over some semiring. We consider four semirings, namely $\mathcal{N}$, the semiring of nonnegative integers; $\mathcal{A}$, the arctical semiring; $\mathcal{T}$, the tropical semiring; and $\mathcal{F}$ the semiring of finite subsets of nonnegative integers.
We show for semirings $\mathcal{N}$ and $\mathcal{A}$ that it is decidable in polynomial time whether or not the cost of accepting computations are bounded, and for $\mathcal{F}$ that it is decidable in polynomial time whether or not the cardinality of occuring cost sets is bounded.
In all these cases we derive explicit upper bounds. For semiring $\mathcal{T}$ we are able to derive similar results at least in case of polynomials of degree at most 1.
For semirings $\mathcal{N}$ and $\mathcal{A}$ the methods are also applicable to get polynomial algorithms for multi-dimensional cost functions.

A grammatical model for Dagstuhl seminars

Franz J. Brandenburg, Universität Passau (joint work with Erzsébet Csuhaj-Varju, Budapest)

We introduce mailbox grammar systems and investigate their computational power. A mailbox grammar system is a grammatical model for the description of cooperating and communicating systems. The model means that some independent agents work on global data and communicate by exchanging messages. The data is manipulated by local replacements in order to achieve a common goal.
Abstractly, a mailbox grammar system consists of n context-free grammers and a network for the exchange of message. For the k-th grammar, a production is a tuple $p = (r, c, E, B, A \rightarrow \alpha, m_1,\ldots,m_n)$, where r is the received message, c is a prefix of r and shall be consumed, E and B are finite sets of existential and blocking context conditions, $A \rightarrow \alpha$ is the context-free core and m_i is the message send to the i-th agent.
We obtain a 1-1 correspondence between Turing machines and mailbox grammar systems. As a consequence, almost all types of mailbox grammar systems are universal., generating the recursively enumerable sets.

On the size of the stack alphabet in PDAs

Detlef Wotschke, Universität Frankfurt (joint work with Jon Goldstine and John Price)

Two transformations are presented which, for any pushdown automaton (PDA) M_o with n_o states and p_o stack symbols, reduce the number of stack symbols to any desired number greater than one. The first transformation preserves deterministic behavior and produces an equivalent PDA with $0(n_o p_o/p)$ states. The second construction, using a technique which does not preserve deterministic behavior, produces an equivalent PDA with $0(n_o \sqrt{p_o}/p)$ states. Both transformations are essentially optimal.

Minimal description length learning

Jorma Rissanen, IBM, San José

On ranking 1-way finitely ambiguous NL languages

Massimiliano Goldwurm, Universität Milano (joint work with Alberto Bertoni)

Given a language $L \subseteq \Sigma^*$, the ranking problem for L consists of computing on input $x \in \Sigma^*$ the number of words $y \in L$ such that either $|y| < |x|$ or $|x| = |y|$ and y is lexicographically lower than x.

We study the problem in the case of 1-way finitely ambiguous NL languages showing that it belongs to DET and hence to the class NC^2. The technique is based on the use of direct sum and Kronecker product for matrices. We also extend this result to a (apparently) larger class of languages including the boolean closure of the class of 1-way finitely ambiguous NL languages.

Indexed grammars - the natural extension of contextfree grammars

Jürgen Duske, Universität Hannover

The talk gives a short overview of some results of indexed grammars which were obtained in the last years at the Institut für Informatik of the University of Hannover.

In 1968, Aho gave the definition of an indexed grammar, which is a slight modification of a context-free grammar. These grammars describe the syntactic structure of their generated languages in a context-free like manner.

We introduce two modes of derivation, the V-mode (call by value) which coincides with the original derivation mode of Aho and the R-mode (call by reference) which allows us to generate all type-0 languages.

Self-embedding indexed grammars are considered, Chomsky-Schützenberger theorems for indexed and type-0 languages are given and indexed LL(k)- and strong indexed LL(k)-grammars are explained. In particular, we have that deterministic context-free languages are exactly the right-linear indexed LL(1)-languages.

As a generalization of the iterated counter automata indexed counter languages has a lot of interesting properties, for example, each type-0 language can be generated by an indexed counter grammar with the R-mode of derivation, that is each type-0 language can be obtained by counting on leftmost derivations of context-free grammars.

Dagstuhl-Seminar 9142

Participants

Jürgen **Albert**
Lehrstuhl für Informatik II
Universität Würzburg
Am Hubland
W-8700 Würzburg
Germany
albert@informatik.uni-wuerzburg.dbp.de
tel.: +49-931-888 5027 (5028)

Jean **Berstel**
Université Paris VI
LITP
Institut Blaise Pascal
4 Place Jussieu
75252 Paris Cedex 05
France
berstel@litp.ibp.fr
tel.: +33-1-44 77 62 42

Franz-Josef **Brandenburg**
Universität Passau
Lehrstuhl für Theoretische Informatik
Innstraße 33
W-8390 Passau
Germany
brandenb@fmi.uni-passau.de
tel.: +49-851-509-343

Anne **Brüggemann-Klein**
Universität Freiburg
Institut für Informatik
Rheinstraße 10-12
W-7800 Freiburg
Germany
brueggemann@informatik.uni-freiburg.dbp.de
tel.: +49-761-203 3897

Didier **Caucal**
Université de Rennes
IRISA
Campus de Beaulieu
Avenue du Général Leclerc
35042 Rennes Cedex
France
caucal@irisa.fr

Bruno **Courcelle**
Université Bordeaux I
Département d'Informatique
351 Cours de la Libération
33405 Talence
France
courcell@geocub.greco-prog.fr
tel.: +33-56 84 60 86

Carsten **Damm**
Humboldt Universität
Fachbereich Informatik
Unter den Linden 6
O-1086 Berlin
Germany
damm@hubinf.uucp

and until 3/92:

TU München
Institut für Informatik
Arcisstarße 21
W-8000 München 2
Germany
damm@informatik.tu-muenchen.de
tel.: +49-89-2105 2397

Jürgen **Dassow**
TU Magdeburg
FG Theoretische Informatik
Postfach 4120
O-3010 Magdeburg
Germany
tel.: +37-91-5923538 / 5922759

Jürgen **Duske**
TU Hannover
Institut für Informatik
Welfengarten 1
W-3000 Hannover 1
Germany
tel.: +49-511-762 5184

Joaquim **Gabarro**
Universidad Politècnica de Catalunya
Facultat d'Informatica
Pau Gargallo 5
E-08028 Barcelona
Spain
gabarro@lsi.upc.es
tel.: +34-3-4017011

Massimiliano **Goldwurm**
Universita degli Studi di Milano
Dip. Scienze dell' Informazione
via Comelico 39
20135 Milano
Italy
goldwurm@imiucca.csi.unimi.it
tel.: +39-2-55 00 63 05

Jozef **Gruska**
Slovak Academy of Sciences
Department of Informatics
Computing Center
Dubravaska cesta 9
842 35 Bratislava
CSFR
tel.: +42-7-374422
and
Universität Hamburg
Fachbereich Informatik
Rothenbaumchaussee 67/69
W-Hamburg 13
gruska@rosun1.informatik.uni-hamburg.de
el: +49-40-4123 5684 (5672)

Günter **Hotz**
Universität des Saarlandes
Fachbereich 14 - Informatik
Im Stadtwald 15
W-6600 Saarbrücken 11
Germany
tel.: +49-681-302 2414

Dung T. **Huynh**
The University of Texas at Dallas
Computer Science Department
Box 830688
Richardson TX 75083
USA
huynh@utdallas.edu
tel.: +1-214-690 2169

Birgit **Jenner**
TU München
Institut für Informatik
Arcisstr. 21
W-8000 München 2
Germany
jenner@informatik.tu-muenchen.de
tel.: +49-89-2105-2387

Helmut **Jürgensen**
University of Western Ontario
Department of Computer Science
Middlesex College
London Ontario N6A 5B7
Canada
helmut@uwo.ca
tel.: +1-519-661-3560

Alica **Kelemenova**
Slovak Academy of Sciences
Department of Informatics
Computing Center
Dubravaska cesta 9
842 35 Bratislava
CSFR
tel.: +42-7-375859

Werner **Kuich**
Technische Universität Wien
Abteilung für Theoretische Informatik
Wiedner Hauptstr. 8-10
A-1040 Wien
Austria
kuich@btx.uucp
tel.: +43-1222-58801 5450

Klaus-Jörn **Lange**
TU München
Institut für Informatik
Arcisstraße 21
W-8000 München 2
Germany
lange@informatik.tu-muenchen.de
tel.: +49-89-2105-2403 (Secr.: 2404)

Hing **Leung**
New Mexico State University
Department of Computer Science
Las Cruces NM 88003
USA
hleung@nmsu.edu
tel.: +1-505-646 1038
and
Universität Frankfurt
FB Informatik (20)
c/o Prof. Wotschke
Postfach 111932
W-6000 Frankfurt 11
Germany
leung@psc.informatik.uni-frankfurt.de
tel.: +49-69-798 3976

Jorma **Rissanen**
IBM Almadan Research Center
K 52/802
650 Harry Road
San Jose CA 95120-6099
USA
rissanen@ibm.com

Jürgen **Schäfer**
Hoechst AG
Aus- und Weiterbildung
Computer Lernzentrum
W-6230 Frankfurt
Germany
tel.: +49-69-305 733

Helmut **Seidl**
Universität Saarbrücken
Fachbereich 14 - Informatik
Im Stadtwald 15
W-6600 Saarbrücken 11
Germany
seidl@sol.cs.uni-sb.de
tel.: +49-681-302 2454

Ray J. **Solomonoff**
Universität des Saarlandes
Fachbereich 14 - Informatik
Im Stadtwald 15
W-6600 Saarbrücken 11
Germany
infi@cs.uni-sb.de
tel.: +49-681-302 6389

Wolfgang **Thomas**
Universität Kiel
Institut für Informatik
Olshausenstraße 40
W-2300 Kiel 1
Germany
wt@informatik.uni-kiel.dbp.de
tel.: +49-431-880 4480

Paul **Vitanyi**
CWI - Mathematisch Centrum
Kruislaan 413
NL-1098 SJ Amsterdam
The Netherlands
paulv@cwi.nl
tel.: +31-20-592 4124

Dietmar **Wätjen**
TU Braunschweig
Institut für Theoretische Informatik
Gaußstr. 11
W-3300 Braunschweig
Germany
waetjen@infbs.uucp
tel.: +49-531-3912387

Hermann **Walter**
Technische Hochschule Darmstadt
Institut für Theoretische Informatik
Frankfurter Str. 69 A
W-6100 Darmstadt
Germany
xitimskr@ddathd21.bitnet
tel.: +49-6151-16 34 05/16 48 95

Egon **Wanke**
Universität GH Paderborn
FB 17 - Mathematik/Informatik
Warburgerstr. 100
W-4790 Paderborn
Germany
egon@uni-paderborn.de
tel.: +49-5251-60 30 74

Andreas **Weber**
Universität Frankfurt
Fachbereich Informatik (20)
Robert-Mayer-Str. 11-15
W-6000 Frankfurt 11
Germany
weber@psc.informatik.uni-frankfurt.de
tel.: +49-69-798 8176

Detlef **Wotschke**
Universität Frankfurt
Fachbereich Informatik (20)
Robert-Mayer-Str. 11-15
W-6000 Frankfurt 11
Germany
wotschke@psc.informatik.uni-frankfurt.de
tel.: +49-69-798 3800

Bisher erschienene und geplante Titel:

W. Gentzsch, W.J. Paul (editors):
Architecture and Performance, Dagstuhl-Seminar-Report; 1, 18.-20.6.1990; (9025)

K. Harbusch, W. Wahlster (editors):
Tree Adjoining Grammars, 1st. International Worshop on TAGs: Formal Theory and Applications, Dagstuhl-Seminar-Report; 2, 15.-17.8.1990 (9033)

Ch. Hankin, R. Wilhelm (editors):
Functional Languages: Optimization for Parallelism, Dagstuhl-Seminar-Report; 3, 3.-7.9.1990 (9036)

H. Alt, E. Welzl (editors):
Algorithmic Geometry, Dagstuhl-Seminar-Report; 4, 8.-12.10.1990 (9041)

J. Berstel , J.E. Pin, W. Thomas (editors):
Automata Theory and Applications in Logic and Complexity, Dagstuhl-Seminar-Report; 5, 14.-18.1.1991 (9103)

B. Becker, Ch. Meinel (editors):
Entwerfen, Prüfen, Testen, Dagstuhl-Seminar-Report; 6, 18.-22.2.1991 (9108)

J. P. Finance, S. Jähnichen, J. Loeckx, M. Wirsing (editors):
Logical Theory for Program Construction, Dagstuhl-Seminar-Report; 7, 25.2.-1.3.1991 (9109)

E. W. Mayr, F. Meyer auf der Heide (editors):
Parallel and Distributed Algorithms, Dagstuhl-Seminar-Report; 8, 4.-8.3.1991 (9110)

M. Broy, P. Deussen, E.-R. Olderog, W.P. de Roever (editors):
Concurrent Systems: Semantics, Specification, and Synthesis, Dagstuhl-Seminar-Report; 9, 11.-15.3.1991 (9111)

K. Apt, K. Indermark, M. Rodriguez-Artalejo (editors):
Integration of Functional and Logic Programming, Dagstuhl-Seminar-Report; 10, 18.-22.3.1991 (9112)

E. Novak, J. Traub, H. Wozniakowski (editors):
Algorithms and Complexity for Continuous Problems, Dagstuhl-Seminar-Report; 11, 15-19.4.1991 (9116)

B. Nebel, C. Peltason, K. v. Luck (editors):
Terminological Logics, Dagstuhl-Seminar-Report; 12, 6.5.-18.5.1991 (9119)

R. Giegerich, S. Graham (editors):
Code Generation - Concepts, Tools, Techniques, Dagstuhl-Seminar-Report; 13, 20.-24.5.1991 (9121)

M. Karpinski, M. Luby, U. Vazirani (editors):
Randomized Algorithms, Dagstuhl-Seminar-Report; 14, 10.-14.6.1991 (9124)

J. Ch. Freytag, D. Maier, G. Vossen (editors):
Query Processing in Object-Oriented, Complex-Object and Nested Relation Databases, Dagstuhl-Seminar-Report; 15, 17.-21.6.1991 (9125)

M. Droste, Y. Gurevich (editors):
Semantics of Programming Languages and Model Theory, Dagstuhl-Seminar-Report; 16, 24.-28.6.1991 (9126)

G. Farin, H. Hagen, H. Noltemeier (editors):
Geometric Modelling, Dagstuhl-Seminar-Report; 17, 1.-5.7.1991 (9127)

A. Karshmer, J. Nehmer (editors):
Operating Systems of the 1990s, Dagstuhl-Seminar-Report; 18, 8.-12.7.1991 (9128)

H. Hagen, H. Müller, G.M. Nielson (editors):
Scientific Visualization, Dagstuhl-Seminar-Report; 19, 26.8.-30.8.91 (9135)

T. Lengauer, R. Möhring, B. Preas (editors):
Theory and Practice of Physical Design of VLSI Systems, Dagstuhl-Seminar-Report; 20, 2.9.-6.9.91 (9136)

F. Bancilhon, P. Lockemann, D. Tsichritzis (editors):
Directions of Future Database Research, Dagstuhl-Seminar-Report; 21, 9.9.-12.9.91 (9137)

H. Alt , B. Chazelle, E. Welzl (editors):
Computational Geometry, Dagstuhl-Seminar-Report; 22, 07.10.-11.10.91 (9141)

F.J. Brandenburg , J. Berstel, D. Wotschke (editors):
Trends and Applications in Formal Language Theory, Dagstuhl-Seminar-Report; 23, 14.10.-18.10.91 (9142)

H. Comon , H. Ganzinger, C. Kirchner, H. Kirchner, J.-L. Lassez , G. Smolka (editors):
Theorem Proving and Logic Programming with Constraints, Dagstuhl-Seminar-Report; 24, 21.10.-25.10.91 (9143)

H. Noltemeier, T. Ottmann, D. Wood (editors):
Data Structures, Dagstuhl-Seminar-Report; 25, 4.11.-8.11.91 (9145)

A. Dress, M. Karpinski, M. Singer(editors):
Efficient Interpolation Algorithms, Dagstuhl-Seminar-Report; 26, 2.-6.12.91 (9149)

B. Buchberger, J. Davenport, F. Schwarz (editors):
Algorithms of Computeralgebra, Dagstuhl-Seminar-Report; 27, 16..-20.12.91 (9151)

K. Compton, J.E. Pin , W. Thomas (editors): Automata Theory: Infinite Computations, Dagstuhl-Seminar-Report; 28, 6.-10.1.92 (9202)

H. Langmaack, E. Neuhold, M. Paul (editors): Software Construction - Foundation and Application, Dagstuhl-Seminar-Report; 29, 13..-17.1.92 (9203)

K. Ambos-Spies, S. Homer, U. Schöning (editors): Structure and Complexity Theory, Dagstuhl-Seminar-Report; 30, 3.-7.02.92 (9206)

B. Booß, W. Coy, J.-M. Pflüger (editors): Limits of Modelling with Programmed Machines, Dagstuhl-Seminar-Report; 31, 10.-14.2.92 (9207)