Advances on Two-dimensional Language Theory
Scientific Report

M. Anselmo [*] *C. Choffrut* [†] *D. Giammarresi* [‡] *J. Kari* [§]

1 Executive summary

The ESF Scientific Meeting "Advances on Two-dimensional Language Theory " was held in Salerno (Italy) from May 3 to May 5, 2006. The Meeting took place in the magnificent Salerno's City Hall (*Sala della Provincia*), kindly offered by the administration of the *Provincia di Salerno*. There were 36 participants from 7 different countries (Austria, Czech Republic, Finland, France, Germany, Italy and USA) (see the complete list of participants in Section 5). Among them there were 4 invited speakers and other 11 speakers (see the final programme in Section 4).

The goal of the workshop was to **gather** a number of active researchers in the field of 2D-languages, in order to present recent results, to exchange ideas and to cooperate on several selected topics of the theory.

The workshop was opened by a welcome speech given by a representative of the Province of Salerno. The program was organized in three plenary talk sessions together with two open problems sessions devoted to working groups. Moreover, each talk session started with an invited lecture followed by some contributing talks and short open problems discussions. The last day was also dedicated to the 60-th birthday of the convenor Christian Choffrut.

No registration fees were required. Lunches and coffee breaks were offered to every participant. Travel and accomodation costs of invited speakers and convenors (except the local one) were reimbursed while only accomodation expenses were reimbursed to other speakers. Two young speakers were also fully refunded upon their request.

All participants manifested their appreciation of the scientific program and the overall organization of the event.

A web page containing all information on the workshop is available since February 2006 at the URL: http://www.dia.unisa.it/conferences/W2DL/.

2 Scientific content of the event

About 50 years ago, Minsky suggested that images could be considered as phrases of a two-dimensional (2D) language, whereby extending the formalisms developed in language theory. In this context, the approaches in terms of grammars (Rosenfeld, 1979) and automata (Blum, Hewitt, 1967) are meaningful. Taking inspiration from ideas and combinatorial problems relevant for strings, researchers from several countries tried to extend some results and techniques to 2D structures.

[*]Dip. Informatica ed Applicazioni, Università di Salerno, 84081 Baronissi (Sa), Italy.
[†]LIAFA, Université Paris VII, 75251 Paris CEDEX 05, France.
[‡]Dip. Matematica, Università di Roma "Tor Vergata", via Ricerca Scientifica, 00133 Roma, Italy.
[§]Department of Mathematics, FIN-20014 University of Turku, Finland.

A basic step in this direction was done in 1991 when A. Restivo and D. Giammarresi defined *tiling systems* and the family REC of 2D languages recognizable by a tiling system. REC family is the class nowadays considered as the 2D counterpart of the class of regular languages on strings.

The workshop started with the invited talk of A. Restivo that overviewed tiling recognizable languages. He showed that different variants of the basic definition give rise to the same family REC and analyzed similarities and differences with respect to the one dimensional case. Finally, he discussed some open problems and research perspectives.

An important fact about REC family is that its definition is intrinsically non-deterministic. Defining a proper definition of determinism on 2D languages is still an open problem. *Determinism*, *unambiguity* and *quantitative aspects* in recognizability were the subjects of the talks given by M. Madonia, I. Mäurer and K. Reinhardt, respectively. M. Madonia talked about the family UREC of unambiguous recognizable languages. Contrarily to the 1D case, there are recognizable languages that are inherently ambiguous. Then she proved some necessary condition for UREC and other properties of the family. Her talk ended with some proposals of definitions of determinism inside REC family to be compared with the other definition of determinism for recognizable languages presented by K. Reinhardt. The latter one is independent from the direction in the picture. K. Reinhardt also considered 2D languages expressing counting and connectivity properties on pictures. In I. Mäurer's talk, tiling systems and 2D automata were generalized to a quantitative setting involving weights and picture series were considered. She presented a characterization theorem of projections of rational picture series in terms of weighted picture automata and some weighted logics. Comparison of deterministic and non-deterministic models for 2D languages recognition was also concerned in J. Kari's talk, where he considered recognition by 4-way automata. Such automata are able to recognize surprisingly complicated languages. Some results were presented regarding the special class of square pictures languages.

Complexity aspects of recognizability of 2-dimensional and more generally n-dimensional languages were considered in the invited talk of B. Borchert and in the talk given by V. Lonati. Based on the notions of locality and recognizability for n-dimensional languages, B. Borchert introduced n-dimensionally colorable 1-dimensional languages. A language L is in NP if and only if L is n-dimensionally colorable for some n. An analogous characterization in terms of deterministic n-dimensional colorability holds for P. The addition of one unbounded dimension for coloring leads to a characterization of PSPACE. V. Lonati presented a characterization of the class of unary square languages that are tiling recognizable in terms of computational complexity. A set L of unary square pictures is in REC if and only if the set of binary encodings of sizes of pictures in L is accepted by a 1-tape nondeterministic Turing machine working in 2^n space with 2^n head reversals. Such a class of binary languages lies between NTIME(2^n) and NTIME(4^n).

A first attempt in the literature to extend the theory of formal languages to the theory of 2-dimensional languages was done considering the formalism of grammars. Since then, several definitions of *2D grammars* have been given. Lately two sorts of grammars have been defined, by O. Matz and then by S. Crespi and M. Pradella (Tile Rewriting Grammars), sharing a structural view of pictures. S. Crespi in his invited talk discussed these models. He showed that when restricted to 1D, both models coincide with type 2 grammars, thus allowing the generation of typical context-free, non-regular languages such as palindromes. But in 2D, Matz's grammars are more restricted than TRG ones. REC family exactly coincides with languages generated by non-recursive TRG rules. The problem of parsing 2D languages was afforded by M. Pradella. In his talk he proposed a polynomial-time parsing approach for languages defined by Matz's grammars which follows the classical Cocke, Kasami and Younger's technique and a novel practical approach for parsing REC languages using recent verification techniques. Two-dimensional array grammars and array automata were considered in the talks given by R. Freund and by M. Oswald. While M. Oswald considered k-level parallel array grammars, also in their more complex variant with attributes, R. Freund considered 2D array grammars as an useful tool for syntactical pattern recognition, especially for hand-written character recognition.

The *2D pattern matching* problem was afforded also in J. Zdarek's talk. His main contribution

was a general finite automata based approach to modelling of 2D pattern matching problems and two particular methods based on it. The first of these two models describes the 2D exact pattern matching and it presents an automata based version of the Bird and Baker approach with lower space complexity than the original algorithm. The second is a model of the 2D approximate pattern matching using the two-dimensional Hamming distance.

A more structural and formal point of view was followed in the talk of F. De Carli. She presented some results on the *lattice structure* of local 2D languages showing some algebraic and combinatorial properties of the lattices of local languages for different sizes of the alphabets.

N. Jonoska investigated REC languages that are defined by the so-called *dot systems*. She showed some results concerning transitivity and mixing properties.

J.-P. Reveillès gave an invited talk about 2D languages from the point of view of an expert in *computational geometry*. He overviewed general digital linear objects. Approximation of lines or planes by point lattices using error analysis or intersection schemes sequences produce very similar objects: digital lines and Sturmian words, either in dimension one or two. But these objects can also be obtained from an arithmetical approach giving a very large amount of generalizations. While these new objects show their usefulness in geometrical applications they also certainly possess many algebraic and combinatorial properties.

More details about the content of the talks can be found in Section 6 where the abstracts of all presentations are collected. During the open problems sessions and the working groups, many interesting problems concerning the topics afforded in the talks were pointed out and discussed, sometimes giving rise to some conjectures about possible solutions. A short collection of open problems can be found in Section 7.

3 Assessment of the results and impact of the event on the future direction of the field

The research on 2D languages is active from several decades: ten years ago the state of the art on this field was analyzed and surveyed in a chapter of the *Handbook of Formal Languages*, G.Rozenberg, *et al.* Eds, Vol. III, pag. 215–268. Springer Verlag, 1997. Since then, the research continued in many directions and several results were presented in important conferences and journals. The invited talks given at the Meeting surveyed some of those important directions, while the contributed talks mostly showed some recent results on the topic. The speakers were all highly valued researchers in the field and the participants well motivated and interested in the area. The atmosphere resulted very friendly and active.

As expected, the workshop was very fruitful because it succeeded both in updating the state-of-art of results on formal two-dimensional languages theory and in collecting open questions and problems. After each presentation there were always questions and comments and the open problems sections lasted many hours of joint work with blackboard discussions. All those discussions have provided a better understanding of the main open problems in the area and of the new directions of research in which to investigate. New collaborations among the participants arose in those days and there were established agreements for exchange visits.

Due to the wide appreciation and satisfaction for the work done in the days of the meeting, many participants expressed their desire to promote other initiatives focused on the study of 2D language theory, and namely in repeating the meeting of the 2D languages community.

We expect that the meeting will provide scientific results and publications in the next years. A technical list of open problems formalized during the meeting's discussions is given in Section 7.

4 Final programme of the meeting

The final program of the Meeting was the following:

- **Wednesday, May 3, 2006**

  14:15 Opening of the workshop

  14:30-15:30 Invited talk: Antonio Restivo "An overview on tiling recognizable languages"

  15:30-16:00 Maria Madonia "Unambiguous recognizable two-dimensional languages"

  16:00-16:30 Coffee break

  16:30-17:00 Ina Mäurer "Characterizations of recognizable picture series"

  17:00-17:30 Klaus Reinhardt "Counting and connectivity in pictures"

  17:30-18:00 Violetta Lonati "On the complexity of unary square tiling recognizable languages"

  18:00-19:00 Open problems session

- **Thursday, May 4, 2006**

  9:30-10:30 Invited talk: Stefano Crespi-Reghizzi "Structural properties of certain picture grammars"

  10:30-11:00 Coffee break

  11:00-11:30 Matteo Pradella "On parsing some classes of 2D languages"

  11:30-12:00 Marion Oswald "2-dimensional k-level parallel array grammars"

  12:00-12:30 Rudolf Freund "Syntactical pattern recognition with 2-dimensional array grammars and array automata"

  12:30-13:00 Open problems session

  13:00-14:30 Lunch break

  14:30-15:30 Invited talk: Jean-Pierre Reveillès "General digital linear objects: arithmetical and Sturm approach"

  15:30-16:30 Working groups

  16:30-17:00 Coffee break

  17:00-19:00 Working groups

- **Friday, May 5, 2006**

  9:30-10:30 Invited talk: Bernd Borchert "Characterizations of P, NP and PSPACE in terms of n-dimensional languages"

  10:30-11:00 Coffee break

  11:00-11:30 Natasha Jonoska "Dot systems and transitivity in two-dimensional languages"

  11:30-12:00 Jan Zdarek "Automata and 2D pattern matching"

  12:00-12:30 Francesca De Carli "A lattice structure of local languages"

  12:30-13:00 Open problems session

  13:00-14:30 Lunch break

  14:30-15:00 Jarkko Kari "4-way finite automata recognizing families of rectangles"

  15:00-16:30 Working groups

  16:30-17:00 Coffee break

  17:00-18:30 Working groups

  18:30- 19:00 Closing of the workshop.

5 List of participants

The participants were 36 (including convenors) and they were from 7 different countries: Austria, Czech Republic, Finland, France, Germany, Italy and USA. They are all listed below.

1. Marcella Anselmo, Università di Salerno, Italy [anselmo@dia.unisa.it]
2. Alberto Bertoni, Università Statale di Milano, Italy [bertoni@dsi.unimi.it]
3. Bernd Borchert, Universität Tübingen, Germany [borchert@informatik.uni-tuebingen.de]
4. Fabio Burderi, Università di Palermo, Italy [burderi@math.math.unipa.it]
5. Alessandra Cherubini, Politecnico di Milano, Italy [alessandra.cherubini@polimi.it]
6. Christian Choffrut, LIAFA, Université Paris VII, France [Christian.Choffrut@liafa.jussieu.fr]
7. Stefano Crespi Reghizzi, Politecnico di Milano, Italy [stefano.crespireghizzi@polimi.it]
8. Flavio D'Alessandro, Università di Roma "La Sapienza", Italy [dalessan@mat.uniroma1.it]
9. Francesca De Carli, Università di Siena, Italy [decarli2@unisi.it]
10. Clelia De Felice, Università di Salerno, Italy [defelice@dia.unisa.it]
11. Aldo De Luca, Università di Napoli, Italy [aldo.deluca@unina.it]
12. Gabriele Fici, Università di Salerno, Italy [fici@dia.unisa.it]
13. Rudolf Freund, Technische Universität, Wien, Austria [rudi@logic.tuwien.ac.at]
14. Andrea Frosini, Università di Siena, Italy [frosini@unisi.it]
15. Dora Giammarresi, Università di Roma "Tor Vergata", Italy [giammarr@mat.uniroma2.it]
16. Massimiliano Goldwurm, Università Statale di Milano, Italy [goldwurm@dsi.unimi.it]
17. Natasha Jonoska, University of South Florida, USA [jonoska@math.usf.edu]
18. Juhani Karhumäki, University of Turku, Finland [karhumak@cs.utu.fi]
19. Jarkko Kari, University of Turku, Finland [jjkari@cs.uiowa.edu]
20. Violetta Lonati, Università Statale di Milano, Italy [lonati@dsi.unimi.it]
21. Maria Madonia, Università di Catania, Italy [madonia@dmi.unict.it]
22. Ina Mäurer, Universität Leipzig, Germany [maeurer@informatik.uni-leipzig.de]
23. Paolo Massazza, Università dell'Insubria, Varese, Italy [paolo.massazza@uninsubria.it]
24. Marion Oswald, Technische Universität, Wien, Austria [marion@logic.at]
25. Joni Pirnot, Manatee Community College, Bradenton, USA [pirnotj@mccfl.edu]
26. Matteo Pradella, CNR-IEIIT c/o Politecnico di Milano, Italy [pradella@elet.polimi.it]
27. Roberto Radicioni, Università Statale di Milano, Italy [radicioni@dsi.unimi.it]
28. Klaus Reinhardt, Universität Tübingen, Germany [reinhard@informatik.uni-tuebingen.de]
29. Antonio Restivo, Università di Palermo, Italy [restivo@math.math.unipa.it]

30. Jean-Pierre Reveillès, Université Clermont 1, France [reveil@llaic3.u-clermont1.fr]
31. Simone Rinaldi, Università di Siena, Italy [rinaldi@unisi.it]
32. Pierluigi Sanpietro, Politecnico di Milano, Italy [sanpietr@elet.polimi.it]
33. Roberto Vaglica, Università di Palermo, Italy [vaglica@math.unipa.it]
34. Stefano Varricchio, Università di Roma "Tor Vergata", Italy [varricch@mat.uniroma2.it]
35. Jan Zdarek, Czech Technical University in Prague, Czech Republic [zdarekj@fel.cvut.cz]
36. Rosalba Zizza, Università di Salerno, Italy [zizza@dia.unisa.it]

6 Abstracts of presentations

6.1 Invited talks

Lecturer: *Bernd Borchert. Universität Tübingen, Germany*
Title: *Characterizations of P, NP, and PSPACE in terms of n-dimensional Languages*
Abstract: Based on the notions of locality and recognizability for n-dimensional languages n-dimensionally colorable 1-dimensional languages are introduced. It is shown: A language L is in NP if and only if L is n-dimensionally colorable for some n. An analogous characterization in terms of deterministic n-dimensional colorability is obtained for P. The addition of one unbounded dimension for coloring leads to a characterization of PSPACE.

Lecturer: *Stefano Crespi Reghizzi. Politecnico di Milano, Italy*
Title: *Structural Properties of Certain Picture Grammars*
Abstract: Several grammar models have been proposed in the course of time to define sets of digital pictures. Such models are often difficult to compare because they rely on rather disparate approaches, although all qualify as extensions in two dimensions of Chomsky's type 2 grammars. Lately two sorts of grammars have been defined, by Matz and then by this author and Pradella (Tile Rewriting Grammars), sharing a structural view of pictures. As in 1D a string is decomposed into constituents, mapped in a syntax tree, now a picture is partitioned into constituent subpictures. The legal layouts of constituents are differently specified, though, in the two models. Matz's grammar use 2D regular expressions, based on vertical and horizontal composition and their closures. TRG grammars use sets of tiles to specify the layout.
Decomposition into constituents is of course recursive. When restricted to 1D, both models coincide with type 2 grammars, thus allowing the generation of typical context-free, non-regular languages such as palindromes.
But in 2D Matz's grammars are more restricted than TRG ones. This can be explained by the inability of 2D regular expressions to generate all pictures made by "pasting" together pictures belonging to a given finite language. Such pasting operation has been formalized by Simplot, and is a natural and powerful generalization of Kleene's star in 2D, as proved by the property that the Tiling Systems of Giammarresi and Restivo are closed under pasting.
A TRG rule can be considered as a way of specifying the structure of a picture, as a partition into unboundedly many homogeneous non-terminal rectangles, which in turn can be similarly decomposed; until at last terminal pixels are plugged in.
We have recently investigated how this structural view of TRG grammars preserves some classical relations between type 3 and type 2 grammars. TS languages exactly coincide with languages generated by non-recursive TRG rules.
With the aim of finding the 2D analogue of type 3 (left-linear or right linear) grammars, we

have defined the notion of corner recursion. Corner recursive TRG grammars always generate TS languages. Both properties, of being non-recursive or corner recursive, are undecidable.

Lecturer: *Antonio Restivo Università di Palermo, Italy*
Title: *An overview on Tiling Recognizable Languages*
Abstract: We introduce tiling recognizable languages (the family REC), considered as the natural generalization to two dimensions of word languages recognized by finite state automata. We first show that different variants of the basic definition give rise to the same family REC. Several characterizations of this family are then presented and, in particular, those in terms of regular expressions and logic formulas. All these results stress the robustness of the notion of tiling recognizability. Similarities and differences with respect to the one dimensional case are analysed. Some interesting links between word languages and languages in REC are also considered. Finally, some open problems and research perspectives are discussed.

Lecturer: *Jean-Pierre Reveillès Université Clermont 1, France*
Title: *General digital linear objects: arithmetical and Sturm approach*
Abstract: Approximation of lines or planes by lattice points using error analysis or intersection schemes sequences produce very similar objects : digital lines and sturmian words, either in dimension one or two (from the sturmian point of view). But these objects can also be obtained from an arithmetical approach giving a very large amount of generalizations. While these new objects show their usefulness in geometrical applications they also certainly possess much algebraic and combinatorial properties. An overview of them will be given.

6.2 Contributed talks

Lecturer: *Maria Madonia Università di Catania, Italy*
Title: *Unambiguous Recognizable Two-dimensional Languages*
Abstract: We consider the family UREC of unambiguous recognizable two-dimensional languages. Informally, a picture language belongs to UREC when it admits an unambiguous tiling system, where a tiling system is unambiguous if every picture has a unique counter-image in its corresponding local language. We prove that there are recognizable languages that are inherently ambiguous, that is UREC family is a proper subclass of REC family. The result is obtained by showing a necessary condition for unambiguous recognizable languages. Further UREC family coincides with the class of picture languages defined by unambiguous 2OTA and it strictly contains its deterministic counterpart. Some closure and non-closure properties of UREC are presented. Finally we show that it is undecidable whether a given tiling system is unambiguous.

Lecturer: *Violetta Lonati. Università degli Studi di Milano – Italy*
Title: *On the Complexity of Unary Square Tiling Recognizable Languages*
Abstract: We give a characterization of the class of unary square languages that are tiling recognizable in terms of computational complexity. We show that a set L of unary square pictures is in REC if and only if the set of binary encodings of sizes of pictures in L is accepted by a 1-tape nondeterministic Turing machine working in 2^n space with 2^n head reversals. Such a class of binary languages lies between NTIME(2^n) and NTIME(4^n). This also implies that there exists a unary square language L such that L is not tiling-recognizable, but the binary string language whose elements represent sizes of pictures in L is recognizable in time $O(4^n log n)$ by a 2-tape nondeterministic Turing machine.

Lecturer: *Matteo Pradella. Politecnico di Milano and CNR IEIIT-MI, Italy*
Title: *On Parsing some Classes of 2D Languages*
Abstract: In this work we address the problem of parsing picture (2D) languages, which is

an important step in 2D language research. We believe that the availability of efficient parsing techniques will enable a wide class of applications of this kind of formalisms.

Among 2D formalisms we cite Giammarresi and Restivo's *Tiling Systems* (TS) [1], that define the *REC* class as the correspondent of regular languages for 2D languages. Unfortunately, [4] showed that the parsing problem for REC languages is NP-complete. Some of the most interesting recent attempts to define analogous of CF languages for pictures are Matz's *Context Free Picture Grammars* (CFPG) [2] and *Tile Rewriting Grammars* (TRG) [3] (containing both CFPG and REC languages).

We propose a polynomial-time parsing approach for CFPG languages which follows the classical Cocke, Kasami and Younger's technique [5]. Also, we present a novel practical approach for parsing TS languages using recent verification techniques.

References

[1] Dora Giammarresi and Antonio Restivo. Two-dimensional languages. In Arto Salomaa and Grzegorz Rozenberg, editors, *Handbook of Formal Languages*, volume 3, Beyond Words, pages 215–267. Springer-Verlag, Berlin, 1997.

[2] Oliver Matz. Regular expressions and context-free grammars for picture languages. In *14th Annual Symposium on Theoretical Aspects of Computer Science*, volume 1200 of *Lecture Notes in Computer Science*, pages 283–294, Lübeck, Germany, 27 February–March 1 1997. Springer-Verlag.

[3] Stefano Crespi Reghizzi and Matteo Pradella. Tile Rewriting Grammars and Picture Languages. *Theoretical Computer Science*, 340(2):257–272, 2005.

[4] Kristian Lindgren, Cristopher Moore, and Mats Nordahl. Complexity of two-dimensional patterns. *Journal of Statistical Physics*, 91(5-6):909–951, June 1998.

[5] D. H. Younger. Recognition of context-free languages in time n^3. *Information and Control*, 10(2):189–208, 1967.

Lecturer: *F. De Carli. Università di Siena. Italy.*

Title: *A lattice Structure of Local Languages*

Abstract: The aim of this work is to define a lattice structure on two-dimensional local languages, and exploit its algebraic and combinatorial properties.

Let Loc be the family of two-dimensional local languages (briefly local languages) over a finite alphabet Σ; this family, with the usual relation of set inclusion, $\langle Loc, \subseteq \rangle$, is a partially ordered set. Then it is possible to define two binary operations, $\vee, \wedge$ such that $\langle Loc, \vee, \wedge \rangle$ is a lattice.

For $n \geq 1$, we use the symbol Loc_n to denote (up to isomorphisms) the lattice of local languages on an alphabet of n elements.

We begin with giving a complete description of Loc_1, from which one can see that Loc_1 is distributive. The task of characterizing Loc_2 is more complex, since, though the lattice is obviously finite, the number of its elements is very large. First we show that Loc_2 is not distributive (and this result can be naturally extended to Loc_n, with $n \geq 2$) and not even modular. This result is obtained by embedding both the lattices $\mathfrak{M}_3$ and $\mathfrak{N}_5$ into Loc_2. Moreover it is possible to prove that Loc_2 has complete chains of different lengths.

Many results on this lattice have been conjectured, and successively checked, using a computer program. In particular we have studied atoms and co-atoms of the lattice, join-irreducible elements, and languages generated by a single word.

Finally we have examined the lattice of horizontal dominoes which is a proper sub-lattice of Loc_2, and showed that it is modular but not distributive. In this case we are able to give a complete characterization of atoms and co-atoms.

Lecturer: *Rudolf Freund. Vienna University of Technology.*

Title: *Syntactical Pattern Recognition with 2-Dimensional Array Grammars and Array Automata*

Abstract: In a series of papers, different variants of 2-dimensional array grammars and 2-dimensional array automata were shown to be useful tools for syntactical pattern recognition, especially for very reliable offline hand-written character recognition. After suitable preprocessing procedures (especially thinning and normalization) of the given hand-written character, the lines representing it within a prescribed rectangular array can be analyzed by suitable array grammars (e.g., see [1]) or array automata (see [2]). Due to the features of the analyzed patterns, a bounded

number of lines can be parsed in parallel (e.g., by k heads of a finite automaton, see [2]); different control mechanisms may guide the analysis (see [3]). For obtaining efficient implementations, hybrid approaches using heuristics for overcoming the problem of scattered lines turned out to be most promising (e.g., see [4]).

References

[1] H. Fernau, R. Freund: Bounded parallelism in array grammars used for character recognition. In: P. Perner, P. S.-P. Wang, A. Rosenfeld (Eds.): Advances in Structural and Syntactical Pattern Recognition, Proc. SSPR'96. LNCS 1121, Springer, Berlin, 1996, 40–49.

[2] H. Fernau, R. Freund, M. Holzer: Character recognition with k-head finite array automata. In: A. Amin, D. Dori, P. Pudil, H. Freeman (Eds.): Advances in Pattern Recognition. Proc. SSPR '98. LNCS 1451, Springer, Berlin, 1998, 282–291

[3] H. Fernau, R. Freund, M. Holzer: Regulated array grammars of finite index. II: Syntactic pattern recognition. In: G. Păun, A. Salomaa (Eds.): Grammatical Models of Multi-Agent Systems. Top. in Comp. Math. 8, Gordon and Breach, 1999, 284–296.

[4] R. Freund, M. Neubauer, M. Summerer, S. Gruber, J. Schaffer, R. Swoboda: A hybrid system for the recognition of hand-written characters. In: F. J. Ferri, J. M. Iñesta-Quereda, A. Amin, P. Pudil (Eds.): Advances in Pattern Recognition. Proc. SSPR 2000. LNCS 1876, Springer, Berlin, 2000, 67–76.

Lecturer: *Natasha Jonoska. University of South Florida, USA.*

Title: *Dot Systems and Transitivity in Two-dimensional Languages*

Abstract: We investigates two-dimensional recognizable languages (in the sense of Giammarresi, Restivo) that are defined by the so-called "dot systems" that are special subgroups of $(Z/2Z)^{Z^2}$. The dot shapes that provide directional transitivity or mixing for the related language are investigated. It is shown that languages defined by parallelogram shapes fail to be transitive in the direction of a defining vector and hence fail to be mixing, while certain triangular shapes guarantee that the factor language of the associated dot system will be mixing. Dot systems belong to a class of two-dimensional shift spaces that have a factor language such that every admissable block can be extended to a configuration of the entire plane. For this class of shift spaces we introduce a finite graph (i.e., a finite state automaton) that recognizes two-dimensional local languages, then show that certain transitivity properties may be observed from the structure of the finite graph.

Lecturer: *Jarkko Kari. University of Turku, Finland.*

Title: *4-way Finite Automata Recognizing Families of Rectangles*

Abstract: In this talk we consider 4-way finite automata operating on unmarked rectangles. We give examples of deterministic automata that recognize surprisingly complicated sets, e.g. squares of sizes 11^p for prime numbers p. Using simulations of two-counter machines we show that sets of squares recognized by 4-way DFA can be as sparse as any recursively enumerable set. More precisely, for any recursively enumerable set $\{a_1, a_2, \dots\}$ of positive integers, there exists a 4-way deterministic finite automaton that recognizes squares of sizes $\{b_1, b_2, \dots\}$ where $b_i > a_i$ for all $i = 1, 2, \dots$. We also show that non-deterministic automata are strictly more powerful than deterministic ones, and that non-deterministic automata are not closed under complementation, even in the case of unmarked rectangles. The talk is based on the following two joint papers with Cristopher Moore.

References

[1] J.Kari, C. Moore. New results on alternating and non-deterministic two-dimensional finite automata. *"Proceedings of STACS'2001, 18th Annual Symposium on Theoretical Aspects of Computer Science".* Lecture Notes in computer Science **2010**, 396–406, Springer-Verlag, 2001.

[2] J. Kari, C. Moore. Rectangles and squares recognized by two-dimensional automata. Theory is Forever, Essays Dedicated to Arto Salomaa on the Occasion of His 70th Birthday. Juhani Karhumäki, Hermann A. Maurer, Gheorghe Paun, Grzegorz Rozenberg (Eds.). Lecture Notes in Computer Science **3113**, 134–144, 2004.

Lecturer: *Ina Mäurer. Universität Leipzig, Germany*

Title: *Characterizations of Recognizable Picture Series*

Abstract: In the literature, a variety of formal models to recognize or generate two-dimensional objects, called pictures, have been proposed. Restivo and Giammarresi defined the family REC of

recognizable picture languages (cf. [GR97]). This family is very robust and has been characterized in terms of types of automata, sets of tiles, rational operations or existential monadic second-order (MSO) logic.

A notion of weighted recognizability for picture languages defined by weighted picture automata (WPA) was introduced in [BG05]. The weights are taken from some commutative semiring. The behavior of a weighted picture automaton is a picture series mapping pictures over an alphabet to some semiring. We will define rational operations and projections on such picture series and generalize devices of tiling systems and 2-dimensional on-line tessellation automata to a quantitative setting involving weights.

Recently, Büchi's fundamental theorem on the coincidence of the class of regular word languages with the family of languages definable in monadic second order logic was generalized to weighted finite automata by introducing a logic with weights [DG05]. Here, we will establish a weighted MSO logic for pictures. The semantics of a weighted formula will be a picture series over a commutative semiring. We will prove the following theorem [Mäu05],[Mäu06]. The required notions will be defined in the talk.

Theorem. Let Σ be an alphabet, K a commutative semiring and $S : \Sigma^{++} \to K$ a picture series. The following assertions are equivalent.

1. S is the behavior of a weighted picture automaton.
2. S is the behavior of a weighted 2-dimensional on-line tessellation automaton.
3. S is the projection of a rational picture series.
4. S is the projection of a tile-local series.
5. S is the semantics of some weighted restricted monadic second-order sentence.

These equivalent weighted picture devices define *recognizable* picture series and can be used to model several application examples, e.g. the intensity of light of a picture (interpreting the alphabet as different levels of gray) or the amplitude of a monochrome subpicture of a colored picture. The presented equivalences generalize the main results of [GR97] to a weighted setting.

References

[BG05] S. Bozapalidis and A. Grammatikopoulou. Recognizable picture series. In M. Droste and H. Vogler, editors, *Special issue on Weighted Automata*, presented at WATA 2004, Dresden, Journal of Automata, Languages and Combinatorics, 2005. In print.

[DG05] M. Droste and P. Gastin. Weighted automata and logics. In *32nd ICALP*, volume 3580 of *Lecture Notes in Computer Science*, pages 513–525. Springer, 2005.

[GR97] D. Giammarresi and A. Restivo. Two-dimensional languages. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3, pages 215–267. Springer, Berlin, 1997.

[Mäu05] I. Mäurer. Recognizable and rational picture series. In *Conference on Algebraic Informatics*, pages 141–155. Aristotle University of Thessaloniki Press, 2005.

[Mäu06] I. Mäurer. Weighted picture automata and weighted logics. In B. Durand and W. Thomas, editors, *STACS 2006*, volume 3885 of *Lecture Notes in Computer Science*, pages 313–324. Springer Berlin, 2006.

Lecturer: *Marion Oswald. Vienna University of Technology.*

Title: 2-*dimensional k-level Parallel Array Grammars*

Abstract: For modelling agents and their actions as well as their interactions with the environment we propose 2-dimensional k-level parallel array grammars (as introduced in [2], also see [1]). This theoretical model allows us to describe the status of the objects under consideration at different levels of scaling, from their atomic reactions to their overall behavior. The dynamic processes within the objects and between different objects are accomplished by the application of parallel rules at each level of the array grammars. To implement a more complex system behavior, we use attributes at specific levels of object description which allow for including specific information as well as for complex interactions.

As a specific example we can model Kohonen's self-organizing feature maps (see [4]) whose underlying grid structures are especially well suited for being simulated by 2-dimensional array grammars. Hierarchical variants of Kohonen's model nicely correspond with the k-level parallel array grammars. The dynamic evolution of self-organizing feature maps can be simulated by using suitable attributes in the parallel array productions. Another variant of attributed parallel array

grammars was already shown to be useful as a specification language for various models of neural networks and as a formal tool for proving specific characteristic features of these networks in [3].

References

[1] A. Binder, R. Freund, M. Oswald, Modelling artificial life by attributed eco-array grammars, *Artificial Life and Robotics* **9**, 1 (2005), 63–66.

[2] R. Freund, Multi-level eco-array systems, in: Gh. Păun (ed.), Artificial life: Grammatical Models, Black Sea University Press, Romania (1995), 175–201.

[3] R. Freund, F. Tafill, Modeling Kohonen networks by attributed parallel array systems, *SPIE Proceedings* Vol. 2093 (1994), 161–172.

[4] T. Kohonen, Self-Organization and Associative Memory, Springer Series in Information Sciences **8**, Heidelberg, 1984.

Lecturer: *Klaus Reinhardt. Universität Tübingen, Germany.*

Title: *Counting and Connectivity in Pictures*

Abstract: An important aspect of counting in formal languages is the ability to count letters in a string. This has always been viewed as a benchmark for language recognition formalisms. For example, $\{a^n b^n | n > 0\}$ is the "prototype" of a language which is not regular. In contrast to this, we show (published in [Rei01]) that the language of pictures (two-dimensional words) with the same number of a's and b's (and an appropriate length-width ratio) is a recognizable picture language. Here, we use counters similar to those used in [Für82] and combine it with the idea of a flow. The picture languages considered here are characterized in [GRS94] as existential monadic second order logic over two-dimensional structures. To describe the power of logic, we consider its ability to express representations of a counter. This also means that counting in two-dimensional arrays is definable in existential monadic second-order logic. A more detailed version can be found in Chapter 4.3 in [Rei05].

Furthermore we show that the language of pictures over $\{a, b\}$, in which all occurring b's are connected, is recognizable. This was published in [Rei98] and solves an open problem in [Mat98]. A more detailed version can be found in Chapter 4.4 in [Rei05]. The result is achieved by organizing the b's in a tree while avoiding non-connected cycles which could otherwise not be distinguished locally. The same problem caused by cycles arises when we consider a certain notion of determinism which is independent from the direction in the picture. The inclusion of the deterministic class in the nondeterministic class is not anymore obvious here. We can generalize the construction for avoiding cycles to show that mono-causal deterministically recognizable picture languages are recognizable.

Finally we discuss possible deterministic methods for counting at least for the case that the picture is a square of a power of the 2 size. References and more information can be found at `http://www-fs.informatik.uni-tuebingen.de/~reinhard/piclang` .

Lecturer: *Jan Zdarek. Czech Technical University in Prague,Czech Republic*

Title: *Automata and 2D Pattern Matching*

Abstract: Our research addresses the problem of modelling of 2D pattern matching by classical finite automata. The main contribution is a general finite automata based approach to modelling of two-dimensional pattern matching problems. Recently, we have presented two particular methods based on the generic algorithm. The first of these two models describes the 2D exact pattern matching and it presents an automata based version of the Bird and Baker approach with lower space complexity than the original algorithm. The second is a model of the 2D approximate pattern matching using the two-dimensional Hamming distance.

Beside automata based models we deal with issues in implementing them. In general it is impossible to use a simulation of nondeterministic models and obtain the linear time complexity. Yet, there exist direct construction methods of equivalent deterministic pattern matching automata. Using them, our method of 2D exact pattern matching has asymptotic time complexity $\mathcal{O}(|T|)$, $|T|$ be the size of the text, $T \in \Sigma^{n \times n}$, Σ be a finite alphabet and n is the length of the rows of T. The space complexity is $\mathcal{O}(|P| + n)$, where $|P|$ is the size of the pattern, $P \in \Sigma^{m \times m}$. It

corresponds to the size of the matching automaton and the size of only one row of the text array in which we search for a pattern.

It is possible to simulate pattern matching automata composing the model of 2D approximate pattern matching using the 2D Hamming distance. Using their simulation by the dynamic programming its time complexity is $\mathcal{O}(|T||P|)$, and it could be improved by e.g. methods using the bit-parallelism. The space complexity is $\mathcal{O}(n|P|)$.

Currently we work towards finding of suitable automata models for further 2D edit distances, like R, C, L, and RC (Baeza-Yates and Navarro, 2000).

The main point of our work is that it reuses a great deal of pattern matching automata in a new area of application and we offer a systematic approach for describing two-dimensional pattern matching using different distances.

7 Open problems

Several open problems arose from the interesting discussions after each talk and during the "open problems" sessions. Some of them are listed below.

Closure properties and REC subclasses

- Extension of Simplot-star to picture series
- Closure of the class of unambiguous picture languages under some unambiguous Simplot-star
- Relations between supports of recognizable picture series and recognizable picture languages, cut-languages, characteristic series for certain commutative semi-rings

Ambiguous picture languages

Define a recognizable language be k-ambiguous if one recognizing tiling system admits at most k pre-images in the corresponding local set for every picture of the language. It is possible to show, that the language L_{col-ij} (set of pictures for which there exist i, j such that the i-th column equal to the j-one) is not k-ambiguous for all k. The proof would be similar to the argument in M.Madonia's talk to prove that L_{col-ij} is not unambiguous.

- Is it possible to find languages that are k-ambiguous but not $(k-1)$-ambiguous for each k or to show that UREC=BUREC (Bounded UREC)?

Automata model for 2D languages

- Are 4-way alternating finite automata (4-AFA) over unary alphabet closed under complementation?
- We know that the complements of 4-AFA are included in REC. Is the inclusion strict ?
- Find some (useful) technique to prove that subsets of squares are not accepted by 4-way DFA or 4-way NFA. In particular, are squares whose size is a prime number recognized by 4-DFA or 4-NFA ?
- Are the inclusions 4-DFA $\subseteq$ 4-NFA $\subseteq$ 4-AFA strict when restricted to squares?

Tile Rewriting Grammars (TRG) and Tiling systems recognizability

It is known that TRG definable picture languages form a family which strictly contains REC family.

- Consider picture languages over a unary alphabet. Does the family of TRG unary languages strictly include the family of unary languages in REC?

Remark that if the two above families were to coincide, the well-known properties that Context-free and Regular unary languages coincide would happen to be true also for picture languages.