# Curriculum Vitae

#### Professor Arseny Shur / Mathematics and Computer Science

#### Abstract

I am a professor with twenty years of academic experience, working in Russia and specializing in combinatorics and related areas of discrete mathematics and computer science. I work in Ural State (now Ural Federal) University, which has a very strong and internationally recognized algebraic school. Olga Kharlampovich (CUNY), Mark Sapir (Vanderbilt), Andrei Bulatov (SFU), Andrei Krokhin (Durham), Mikhail Volkov (UrFU), and many more people, including myself, all originated from this school (see, e.g., http://genealogy.math.ndsu.nodak.edu/id.php?id=86245). My current research interests are centered at combinatorics and algorithmics of words (strings), with close connections to graph theory, automata theory, and enumerative combinatorics. I have much experience in teaching at both undergraduate and graduate level, including the design of new courses and textbook writing. I also perform a lot of graduate supervision: six my PhD students have successfully graduated; the number of MSc defenses is above 20.

## Contents

Education, Employment, Awards, Contacts	1
Current Research Interests	2
Publications	3
Overview of Research Contributions	7
Conferences, Workshops, Presentations	9
Active Service to the Academic Community	11
Teaching and Supervision	12
Research/Project Funding Received	14

## Education, Employment, Awards, Contacts

#### **Educational Background**

- 2010 Dr.Sc. in Mathematics (Habilitation), Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Science, Russia. Combinatorial characterization of formal languages
- 1998 Ph.D. in Mathematics, Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Science, Russia. Algebraic and combinatorial properties of equational languages
- 1993 M.Sc. in Mathematics, Ural State University, Russia. Algebraic properties of overlap-free and cube-free words

#### **Employment History at Academic Institutions**

All positions in the table below refer to the Ural State University (Ural Federal University since 2011), Ekaterinburg, Russia.

Sep 2011 - Current	Full Professor, Mathematics and Computer Science, Institute of Mathematics
	and Computer Science
Sep 2001 - Aug 2011	Associate Professor, Mathematics and Computer Science, Faculty of Mathe-
	matics and Mechanics
Sep 1997 - Aug 2001	Assistant Professor, Mathematics, Faculty of Mathematics and Mechanics
Feb 1995 - Aug 1997	Research Officer, Mathematics, Institute of Physics and Applied Mathematics
Sep 1993 - Oct 1994	Researcher, Mathematics, Faculty of Mathematics and Mechanics

### Awards, Honors and Scholarships

Year	Title	Description	Type	Organization
2015	Best Paper Award	at the 41th SOFSEM	Research	41th SOFSEM Conference PC
		Conference		
2013	Professor	Scientific title (hon-	Research,	Ministry of Science and Edu-
		ourable mention)	Teaching	cation of Russia
2012	Best Paper Award	at the 37th MFCS sym-	Research	European Association for The-
		posium		oretical Computer Science
2004	Docent	Scientific title (hon-	Research,	Ministry of Science and Edu-
		ourable mention)	Teaching	cation of Russia
1998	Award to the Best Young		Research	Ural Mathematical Society
	Mathematician			
1993	M.Sc. "Red" Diploma	Highest award for the	Other	Ural State University
		grads during the study		

#### Languages

Russian (mother tongue), English (fluent)

#### Contacts

Address:	Department of Algebra and Discrete Mathematics, Ural Federal University
	620000 pr. Lenina 51, Ekaterinburg, Russia
Phone:	+7 (343) 3899468 (work)
E-mail:	arseny.shur@urfu.ru
Webpage:	http://kadm.imkn.urfu.ru/pages.php?id=shureng

## **Current Research Interests**

- Combinatorics of words
  - Periodicity properties, avoidability relations, growth functions, probabilistic issues related to different types of finite and infinite sequences of symbols; structural theorems and growth issues for sets of such sequences
  - Keywords: words, infinite words, partial words, circular words, power-free words, Abelian power-free words, patterns, squares, cubes, palindromes, periodicity, interaction of periods, avoidability, factor complexity, expectation
- String algorithms
  - Generating repetition-free words, searching and counting palindromes and repetitions
  - Keywords: random repetition-free words, palindromes, squares, runs, Zimin words
- Growth functions of formal languages
  - Classification theorems, asymptotic formulas, and approximation algorithms for calculating growth functions, growth rates etc.

- Keywords: formal languages, combinatorial complexity, growth rate, finite automata
- Graphs and words
  - Non-repetitive colourings, connections between random graphs and random words
  - Keywords: non-repetitive colouring, random graph, random word
- Finite automata
  - Words separation problem, Aho-Corasick automata, probabilistic, quantum and generalized finite automata
  - Keywords: finite automata, forbidden words, automata on real numbers
- Syntactic and relatively free monoids
  - Structural theorems, word problem decidability
  - Keywords: syntactic monoid, word problem, context equivalence

## Publications

### **Refereed Conference Proceedings**

- D. Gasnikov, A. M. Shur. Ternary square-free partial words with many wildcards. Proc. DLT 2016: Springer, 2016. 177-189 (LNCS Vol. 9840)
- P. Gawrychowski, O. Merkurev, A. Shur, and P. Uznanski. Tight Tradeoffs for Real-Time Approximation of Longest Palindromes in Streams. Proc. CPM 2016: Schloss Dagstuhl Leibniz-Zentrum f
  ür Informatik, 2016. 18:1-18:13. (LIPIcs Vol. 54)
- M. Rubinchik, A. M. Shur. EERTREE: An efficient data structure for processing palindromes in strings. Proc. IWOCA 2015: Springer, 2016. 321-333. (LNCS Vol. 9538)
- E. A. Petrova, A. M. Shur. On the tree of ternary square-Free words. Proc. WORDS 2015. Springer, 2015. 223-236. (LNCS Vol. 9304)
- D. Kosolobov, M. Rubinchik, A. M. Shur. Pal<sup>k</sup> is linear recognizable online. Proc. 41th Int. Conf. on Theory and Practice of Computer Science (SOFSEM 2015). Springer, 2015. 289-301. (LNCS Vol. 8939). Best Paper Award
- 26. M. Rubinchik, A. M. Shur. On the number of distinct subpalindromes in words. Proc. 3rd Russian Finnish Symp. on Discrete Mathematics. Inst. Appl. Math. Research, Petrozavodsk, 2014. 96-98.
- A. M. Shur, A. Yakaryilmaz. Quantum, stochastic, and pseudo stochastic languages with few states. Proc. UCNC 2014. Springer, 2014. 327-339. (LNCS Vol. 8553).
- D. Kosolobov, M. Rubinchik, A. M. Shur. Finding distinct subpalindromes online. Proc. Prague Stringology Conf. 2013. Prague: Czech Technical University, 2013. 63-69.
- A. M. Shur. Languages with a finite antidictionary: growth-preserving transformations and available orders of growth. Proc. 17th Int. Conf. on Developments in Language Theory. Berlin: Springer, 2013. 445-457. (LNCS Vol. 7907)
- 22. L. A. Idiatulina, A. M. Shur. Periodic partial words and random bipartite graphs. Proc. 2nd Russian Finnish Symp. on Discrete Mathematics, Turku, 2012. 25-28. (TUCS Lecture Notes 17)
- A. V. Samsonov, A. M. Shur. Binary patterns in binary cube-free words: avoidability and growth. Proc. 14th Mons Days of Theoretical Computer Science. Catholic Univ. Louvain, Louvain-la-Nueve, 2012. 7pp. (Electronic)

- I. N. Tunev, A. M. Shur. On two stronger versions of Dejean's conjecture. Mathematical foundations of computer science. Berlin: Springer, 2012. 800-812. (LNCS Vol. 7464). EATCS Best Paper Award
- E. A. Petrova, A. M. Shur. Constructing premaximal ternary square-free words of any level. Mathematical foundations of computer science. Berlin: Springer, 2012. 752-763. (LNCS Vol. 7464).
- N. Rampersad, J. Shallit, A. Shur. Fife's theorem for (7/3)-powers. Proc. 8th Int. Conf. WORDS 2011. EPTCS Vol.63 (2011), 189-198.
- I. A. Gorbunova, A. M. Shur. On Pansiot words avoiding 3-repetitions. Proc. 8th Int. Conf. WORDS 2011. EPTCS Vol.63 (2011), 138-146.
- E. A. Petrova, A. M. Shur. Constructing premaximal binary cube-free words of any level. Proc. 8th Int. Conf. WORDS 2011. EPTCS Vol.63 (2011), 168-178.
- 15. A. M. Shur. Growth properties of power-free languages. Proc. 15th Int. Conf. on Developments in Language Theory. Berlin: Springer, 2011. 28-43. (LNCS Vol. 6795).
- 14. A. N. Plyushchenko, A. M. Shur. On Brzozovski's conjecture for the free Burnside semigroup satisfying  $x^2 = x^3$ . Proc. 15th Int. Conf. on Developments in Language Theory. Berlin: Springer, 2011. 362-373. (LNCS Vol. 6795).
- K. M. Likhomanov, A. M. Shur. Two combinatorial criteria for BWT images. Proc. 6th International Computer Science Symposium in Russia. Berlin: Springer, 2011. 385-396. (LNCS Vol. 6651).
- A. V. Samsonov, A. M. Shur. On Abelian repetition threshold. Proc. 13th Mons Days of Theoretical Computer Science. Univ. de Picardie Jules Verne, Amiens, 2010. 1–11.
- 11. A. M. Shur. On the existence of minimal  $\beta$ -powers. Proc. 14th Int. Conf. on Developments in Language Theory. Berlin: Springer, 2010. 411-422. (LNCS Vol. 6224).
- 10. A. M. Shur. Growth of power-free languages over large alphabets. Proc. 5th International Computer Science Symposium in Russia. Berlin: Springer, 2010. 350-361. (LNCS Vol. 6072).
- 9. A. M. Shur. Two-sided bounds for the growth rates of power-free languages. Proc. 13th Int. Conf. on Developments in Language Theory. Berlin: Springer, 2009. 466-477. (LNCS Vol. 5583).
- 8. A. M. Shur. Combinatorial complexity of regular languages. Proc. 3rd International Computer Science Symposium in Russia. Berlin: Springer, 2008. 289-301. (LNCS Vol. 5010).
- A. M. Shur, I. A. Gorbunova. On the growth rates of complexity of threshold languages. In: 12th Mons Days of Theoretical Computer Science, Int. Proc. Univ. de Mons-Hainaut, Mons, 2008, 10pp.
- 6. A. N. Plyushchenko, A. M. Shur. Almost overlap-free words and the word problem for the free Burnside semigroup satisfying  $x^2 = x^3$ . Proc. 6th Int. Conf. WORDS, 2007, 245-253, Institut de Mathématiques de Luminy, Marseille, France, 2007.
- 5. A. M. Shur. Factorial languages of low combinatorial complexity. Proc. 10th Int. Conf. on Developments in Language Theory. Berlin: Springer, 2006. 397-407. (LNCS Vol. 4036).
- 4. A. M. Shur. Comparing complexity functions of a language and its extendable part. In: 11th Mons Days of Theoretical Computer Science, Int. Proc. 323-327, IRISA-Rennes, Univ. Rennes, Rennes, 2006.
- A. M. Shur, Y. V. Gamzova. Periods' interaction property for partial words. Proc. 4th Int. Conf. WORDS, Turku, 2003. 75-82. (TUCS Gen. Publ., Vol. 27).

- 2. A. M. Shur, Y. V. Konovalova. On the periods of partial words. Proc. 26th Int. Symp. Mathematical Foundations of Computer Science. Berlin: Springer, 2001. 657-665. (LNCS Vol. 2136).
- E. V. Sukhanov, A. M. Shur. Galois connection in avoidability theory. In: S. Kublanovsky, A. Mikhalev, J. Ponizovsky, P. Higgins eds., Semigroups with applications, including semigroup rings. International conference in honour of E.S. Lyapin, St.-Petersburg, 1995. St.-Petersburg: Severny Ochag, 1999, 397-400.

### **Refereed Journal Articles**

- C. Guo, J. Shallit, A. M. Shur. Palindromic rich words and run-length encodings. Inform. Process. Lett. Vol. 116(12) (2016), 735-738
- M. Rubinchik, A. M. Shur. The number of distinct subpalindromes in random words. Fundamenta Informaticae Vol. 145(3) (2016), 371-384
- 34. A. M. Shur, A. Yakaryilmaz. More on quantum, stochastic, and pseudo stochastic languages with few states. Natural Computing Vol. 15(1) (2016), 129-141
- 33. C. F. Du, J. Shallit, A. M. Shur. Optimal bounds for the similarity density of the Thue-Morse word with overlap-free and 7/3-power-free infinite binary words. Internat. J. Foundat. Comput. Sci. Vol. 26(8) (2015), 1147-1166
- A. M. Shur. Generating square-free words efficiently. Theoretical Computer Science Vol. 601 (2015), 67-72
- W. Rytter, A. M. Shur. Searching for Zimin patterns. Theoretical Computer Science Vol. 571 (2015), 50-57
- A. M. Shur. Languages with a finite antidictionary: some growth questions. Internat J. Foundations of Comput. Sci. Vol. 25(8) (2014), 937-954.
- R. Mercaş, P. Ochem, A. V. Samsonov, A. M. Shur. Binary patterns in binary cube-free words: Avoidability and growth. RAIRO - Theoretical Informatics and Applications Vol. 48(4) (2014), 369-389.
- L. A. Idiatulina, A. M. Shur. Periodic partial words and random bipartite graphs. Fundamenta Informaticae Vol. 132(1) (2014), 15-31.
- A. M. Shur. Growth of power-free languages over large alphabets. Theory of Computing Systems Vol. 54(2) (2014), 224-243.
- I. A. Gorbunova, A. M. Shur. On Pansiot words avoiding 3-repetitions. Internat. J. Foundations of Comp. Sci. Vol. 23(8) (2012), 1583-1594.
- E. A. Petrova, A. M. Shur. Constructing premaximal binary cube-free words of any level. Internat. J. Foundations of Comp. Sci. Vol. 23(8) (2012), 1595-1610.
- A. M. Shur. Growth properties of power-free languages. Computer Science Review Vol. 6 (2012), 187-208.
- A. M. Shur. Deciding context equivalence of binary overlap-free words in linear time. Semigroup Forum Vol.84(3) (2012), 447-471.
- A. V. Samsonov, A. M. Shur. On Abelian repetition threshold. RAIRO Inform. Theor. Appl. Vol. 46 (2012), 147-163.
- A. M. Shur. On the existence of minimal β-powers. Internat. J. Foundations of Comp. Sci., Vol. 22(7) (2011), 1683-1696.

- 20. A. N. Plyushchenko, A. M. Shur. Almost overlap-free words and the word problem for the free Burnside semigroup satisfying  $x^2 = x^3$ . Int. J. Algebra and Comput. Vol. 21(6) (2011), 973-1006.
- A. M. Shur. On ternary square-free circular words. Electronic J. Combinatorics Vol. 17(1) (2010). Article R140.
- A. M. Shur. Growth rates of complexity of power-free languages. Theor. Comput. Sci. Vol. 411 (2010), 3209-3223.
- 17. A. M. Shur. Growth of power-free languages: Numerical and asymptotic bounds. Dokl. Math. Vol. 81(3) (2010), 406-409.
- A. M. Shur. Calculating parameters and types of growth of combinatorial complexity of regular languages. Proc. Inst. Mathematics and Mechanics of Ural Branch of Russian Academy of Sciences Vol. 16(2) (2010), 270-287. In Russian, with English abstract.
- A. M. Shur, I. A. Gorbunova. On the growth rates of complexity of threshold languages. RAIRO Inform. Theor. Appl. Vol. 44 (2010), 175-192.
- 14. A. M. Shur. Languages with finite antidictionaries: growth index and properties of automata. Proc. Ural State Univ. Vol. 74 (2010), 220-245. In Russian, with English abstract.
- A. M. Shur. Growth rates of power-free languages. Russian Mathematics (Iz VUZ) Vol. 53(9) (2009), 73-78.
- A. M. Shur. Polynomial languages with finite antidictionaries. RAIRO Theor. Inform. and Appl. Vol. 43(2) (2009), 269-280.
- A. M. Shur. On intermediate factorial languages. Discrete Applied Mathematics Vol. 157(7) (2009), 1669-1675.
- A. M. Shur. Comparing complexity functions of a language and its extendable part. RAIRO Theor. Inf. Appl. Vol. 42 (2008), 647-655.
- A. M. Shur. Rational approximations of polynomial factorial languages. Int. J. Foundat. Comp. Sci. Vol. 18 (2007), 655-665.
- A. M. Shur. Combinatorial complexity of rational languages. Diskretn. Anal. Issled. Oper. Ser. 1 Vol. 12(2) (2005), 78-99. In Russian.
- A. M. Shur, Y. V. Gamzova. Partial words and the interaction property of periods. Izv. Math. Vol. 68(2) (2004), 405-428.
- A. M. Shur. The structure of the set of cube-free Z-words over a two-letter alphabet. Izv. Math. Vol. 64(4) (2000), 847-871.
- E. V. Sukhanov, A. M. Shur. A class of formal languages. Algebra and Logic. Vol. 37(4) (1998), 270-277.
- A. M. Shur. Syntactic semigroups of avoidable languages. Siberian Math. J. Vol. 39(3) (1998), 594-610.
- 3. A. M. Shur. Binary avoidability and Thue-Morse words. Dokl. Math. Vol. 53(3) (1996), 405-406.
- 2. A. M. Shur. Overlap-free words and Thue-Morse sequences. Int. J. Algebra and Comput. Vol. 6(3) (1996), 353-367.
- 1. A. M. Shur. Binary words avoided by the Thue-Morse sequence. Semigroup Forum. Vol. 53(2) (1996), 212-219.

## **Books and Chapters**

- 5. B. M. Vernikov, A. M. Shur. Graphs and Graph Algorithms. Textbook for students in applied informatics. Ekaterinburg, 2012. 160pp. ISBN 978-5-904440-21-3. In Russian.
- A. P. Zamyatin, A. M. Shur. Languages, Grammars, Acceptors. Textbook for students of mathematical departments. Ural University, Ekaterinburg, 2007. 248pp. ISBN 978-5-7996-0273-0. In Russian.
- 3. A. M. Shur. Combinatorics of Words. Textbook for students of mathematical departments. Ural University, Ekaterinburg, 2003. 96pp. ISBN 978-5-7996-0168-8. In Russian.
- E. V. Sukhanov, A. M. Shur. Problems of entrance exams in mathematics at Ural State University. In: Problems of entrance exams in mathematics at higher education institutions of Ekaterinburg. Ekaterinburg, 1997. 5-14. In Russian.
- V. V. Rasin, D. S. Ananichev, M. V. Volkov, A. G. Gein, S. P. Homenko, V. P. Ivanova, A. M. Shur. Entrance exams in mathematics at Ural State University. Ural University, Ekaterinburg, 2003. 44pp. In Russian.

## **Edited Volumes**

- 4. Special Issue Developments in Language Theory (DLT 2014). Guest Editor: Arseny Shur. International Journal of Foundations of Computer Science Vol. 27(2) (2016).
- Arseny M. Shur, Mikhail V. Volkov (Eds.): Developments in Language Theory 18th International Conference, DLT 2014, Ekaterinburg, Russia, August 26–29, 2014. Proceedings. Springer, 2014. (LNCS Vol. 8633). ISBN 978-3-319-09697-1
- Andrei A. Bulatov, Arseny M. Shur (Eds.): Computer Science Theory and Applications 8th International Computer Science Symposium in Russia, CSR 2013, Ekaterinburg, Russia, June 25– 29, 2013. Proceedings. Berlin: Springer, 2013. (LNCS Vol. 7913). ISBN 978-3-642-38535-3
- Alexander S. Kulikov, Arseny M. Shur (Eds.): Algorithms & Complexity. Abstracts of Reports and Other Materials of the 6th School "Computer Science Days in Ekaterinburg", Ekaterinburg, Russia, June 29 – July 01, 2013. ISBN 978-5-7996-0922-1

## Other Publications

- Combinatorial characterization of formal languages. Doctoral (habilitation) thesis. Institute of Mathematics and Mechanics of Ural Branch of Russian academy of Sciences. Ekaterinburg, Russia. 2010. Manuscript, 287pp. In Russian. English translation of extended abstract is available at: arXiv:1010.5456 [cs.FL]
- 1. Numerical values of the growth rates of power-free languages. arXiv:1009.4415[cs.FL]

## Works Accepted for Publication

1. M. Rubinchik, A. M. Shur. EERTREE: an efficient data structure for processing palindromes in strings, submitted to European J. Combinatorics, 28p.

## Works Submitted for Publication

- 2. J. Kärkkäinen, D. Kempa, Y. Nakashima, S. J. Puglisi, and A. M. Shur. On the size of Lempel-Ziv and Lyndon factorizations, submitted to STACS 2017, 12p.
- 1. A. A. Bulatov, O. Karpova, A. M. Shur, and K. Startsev. Lower bounds on words separation: are there short identities in transformation semigroups?, submitted to Electronic J. Combinatorics, 2016, 12p.

## **Overview of Research Contributions**

Citations refer to the list of publications, with  $\{C=Conference, J=Journal, A=Accepted\}$  papers.

#### **Growth Functions of Formal Languages**

Combinatorial complexity (growth function, census function) of a language  $L \subseteq \Sigma^*$  is the function  $C_L(n) = |L \cap \Sigma^n|$ . It is studied since the seminal papers by Morse and Hedlund (1938, 1940) on symbolic dynamics. We provided a general framework to study this characteristic for different classes of languages.

A classical result of A. Salomaa and Soittola on formal powers series (1978) implies that the combinatorial complexity of any regular language L is asymptotically equivalent to a function  $f_L(n) = p_{n \mod r}(n) \cdot \alpha_{n \mod r}^n$  for some integer r, algebraic numbers  $\alpha_i$ , and polynomials  $p_i$ . We related the parameters of this function to the characteristics of finite automata and provided efficient algorithms computing these characteristics [C8, C23, J8, J14, J16, J18, J30]. Our results can be also used to compute spectral characteristics of arbitrary digraphs.

The objects of most studies on combinatorial complexity are languages closed under taking factors of words (factorial languages). We developed a general approach to the study of combinatorial complexity for factorial languages through regular languages and show the limits of its applicability [C5, J9, J10].

Among factorial languages, the languages avoiding different sorts of repetitions (repetition-free languages) are of the biggest interest from the complexity point of view. We developed efficient algorithms for upper and lower bounds for the growth rates of various types of repetition-free languages, and performed extensive numerical studies for power-free, Abelian-power-free, and pattern-avoiding languages [C8, C9, C12, C15, J15, J18, J22, J29]. Further, we proved analytic asymptotic formulas for the growth rates of power-free languages [C10, J27] and described zeroes of combinatorial complexity for languages of minimal powers [C11, J19, J21]. Overall, these studies, partially joint with my students I. Gorbunova and A. Samsonov, revealed the general picture of growth properties for repetition-free languages, incorporating all previously known results on this topic (see the survey [J24] for more details).

The languages of "rich" words, which are words with the maximum number of distinct palindromes inside, are a good example of interesting factorial languages which are not repetition-free. With C. Guo and J. Shallit we gave the bounds on the growth for binary rich words [J36]; this language seems to have an intermediate (superpolynomial and subexponential) growth, which is a quite rare case.

#### **Repetition Threshold and Related Problems**

Repetition threshold is the smallest number RT(k) such that infinitely many k-ary words contain no  $\beta$ -power with  $\beta > RT(k)$ . The formula for RT(k) is the subject of the famous conjecture by Dejean (1972), proved in part by several authors and finally settled in 2009. A lot of problems arose around this conjecture. The smallest infinite power-free languages, called threshold languages, are of particular interest.

Two types of problems concerning threshold languages are whether the sets of threshold words remain infinite under additional restrictions and whether the combinatorial complexity of threshold languages is exponential. Together with my students I. Gorbunova and I. Tunev, we introduced new methods of study for threshold languages. In [C20], we proved that for a series of alphabets there are exponentially many threshold words satisfying a very restrictive additional condition. In [C17, J15, J26] we proved several structural results giving an evidence that threshold languages not only have the exponential growth for  $k \geq 3$ , but their growth rates tend to a limit  $\approx 1.24$  as k approaches infinity (the corresponding conjecture was formulated in [J15]).

Together with my student A. Samsonov, we studied [C12, J22] an analog of the repetition threshold for Abelian powers, proved lower and upper bounds for it and stated an analog of Dejean's conjecture for Abelian powers.

Together with R. Mercaş, P. Ochem, and A. Samsonov, we classified all binary patterns with respect to the (in)finiteness of the number of avoiding binary cube-free words [C21, J29]. This is one of the very few classification results in the area of pattern avoidance since the mid-1990's.

#### Structure of Repetition-Free Languages

Our contributions here are related to the most important objects in this area: the ternary square-free language SF, the binary cube-free language CF, and the binary overlap-free language OF. The depth of a word w in a factorial language L is the supremum of the lengths of words u such that  $wu \in L$ . The existence of words of arbitrary big finite depth indicates that the structure of L is complicated. Allouche and Shallit asked in 2003 whether there are such words in SF. Together with my student E. Petrova, we answered this question in the affirmative [C19], and give an affirmative answer to the same question about CF [C16, J25]. Moreover, we solved the two-sided versions of both questions as well.

Further structural results about SF were obtained in [C28]: it was proved that the prefix tree of ternary square-free words branches almost uniformly and has no long non-branching path. This is a big improvement over previous results by Shelton and Sony (1981) and Currie (1994,1996). Studying square-free ternary circular words, we found their intimate connection to the paths in the weighted  $K_{3,3}$  graph [J19]. Together with my student D. Gasnikov, we studied square-free ternary partial words and found the exact value of the maximum density of undefined positions in it [C31].

As to the language OF, it has simpler structure (see, e.g., papers by Restivo and Salemi), grows polynomially and is essentially based on a single combinatorial structure: Thue's morphism and its fixed point, called the Thue-Morse word. We provide a linear-time solution to the word problem in the syntactic monoid of OF [J23], which is quite unusual for this type of problems. Together with C. F. Du and J. Shallit we proved that any infinite overlap-free word is surprisingly far from the "pivot" Thue-Morse word in terms of the Hamming distance [J33].

#### **Periodic Partial Words**

A basic property of periodic words is expressed by the Fine-Wilf theorem (1965). The theorem states that a sufficiently long word with periods p and q necessarily has the period gcd(p,q) and gives the sharp bound for what is "sufficiently long". The most interesting generalization of this property, first mentioned by Berstel and Boasson in 1999, is obtained by replacing words with partial words. In this way, one more numerical parameter gets involved, which is the number k of undefined symbols in the word. Together with my student Y. Konovalova (Gamzova) we proved that the analog of the Fine-Wilf property holds for arbitrary partial words, and the minimum length L(k, p, q) of a word, guaranteeing this property, is in direct proportion to k with parameters p, q [C2]. Moreover, we showed that, apart from a few initial values of k, the function L(k, p, q) is the sum of a linear and periodic functions (in k), and describe them both [C3, J7]. In addition, we proved, together with my student L. Idiatulina, a probabilistic version of this result, establishing a phase transition for the Fine-Wilf property [J28].

#### **Algorithmic Problems on Palindromes**

Galil and Seiferas asked in 1978 whether there exists a linear-time algorithm that, given a word w and a positive integer k, decides whether w is a concatenation of k non-empty palindromes. This question was then restated by Crochemore and Rytter in "Jewels of Stringology" (2002). Together with my students D. Kosolobov and M. Rubinchik, we solved this question positively in a stronger form, providing an online linear-time algorithm [C27]. Further, with M. Rubinchik we introduced a new data structure for processing palindromes in strings and solve several palindrome-related problems with its help [C29, A1]. With P. Gawrichovski, P. Uznanski, and my student O. Merkurev we solved the problem of finding the longest palindrome in a string in the streaming model [C30], improving over the results of P. Berebrink et al. (2014).

#### **Properties of Random Words**

We solved some problems related to repetitions in random words. Together with M. Rubinchik, we gave sharp estimates on the expected number of palindromes and squares in a string [J35]. A simple and efficient algorithm for generating square-free words from a random source was proposed and analysed in [J32]. For the Fine-Wilf property of partial words (see above), we proved, together with my student L. Idiatulina, that this property of a random partial word is equivalent to the edge connectedness of a random bipartite graph, and found the phase transition for this graph property in most cases [J28].

## **Conferences**, Workshops, Presentations

### Invited Talks and Workshops

- 2016 Workshop "Combinatorics on Words", CIRM, Marseille, France. "Some Open Problems on Repetitions in Words"
- 2015 Workshop on Combinatorics and Applications at SJTU, Shanghai, China. "The Expected Number of Repetitions in Random Words"
- 2014 15th Mons Days of Theoretical Computer Science (JM 2014). Nancy, France. "Regular approximations of factorial languages: what are they good for?"
- 2014 Combinatorics and algorithmics of strings, Dagstuhl Seminar 14111, Dagstuhl, Germany. "Efficient generation of repetition-free words"
- 2012 2nd Russian Finnish Symposium on Discrete Mathematics (RuFiDiM). Turku, Finland. "Periodic partial words and random bipartite graphs"
- 2011 15th International Conference on Development in Language Theory (DLT 2011). Milan, Italy. "Growth properties of power-free languages"
- 2011 Combinatorial and Algorithmic Aspects of Sequence Processing, Dagstuhl Seminar 11081, Dagstuhl, Germany. "Context equivalence problem for power-free words"
- 2009 International Seminar "Theoretical computer science: from foundations to applications". Niš, Serbia. "Growth properties of power-free languages"
- 2009 International conference on combinatorial algebra. Guangzhou, China. "Growth properties of formal languages"
- 2008 Russian-Indian workshop on Algebra, Combinatorics and Complexity. Ekaterinburg, Russia. "Threshold languages: structure and growth"
- 2008 Russian-Indian workshop on Algebra, Combinatorics and Complexity. Moscow, Russia. "Languages with finite antidictionary"
- 2005 International algebraic conference dedicated to the centennial of the birthday of P.G. Kontorovich and to the 70th birthday of L.N. Shevrin, Ekaterinburg, Russia. "On complexity of formal languages"

### **Conference Talks**

- 2016 20th International Conference on Developments in Language Theory (DLT 2016). Montreal, Canada. "Ternary square-free partial words with many wildcards"
- 2016 27th Annual Symposium on Combinatorial Pattern Matching (CPM 2016). "Tight tradeoffs for real-time approximation of longest palindromes in streams"
- 2014 3rd Russian Finnish Symposium on Discrete Mathematics (RuFiDiM 2014). Petrozavodsk, Russia. "On the number of distinct subpalindromes in words"
- 2014 13th International Symposium on Unconventional Computation and Natural Computation (UCNC 2014). London, ON, Canada. "Quantum, stochastic, and pseudo stochastic languages with few states"
- 2013 17th International Conference on Developments in Language Theory (DLT 2013). Paris, France. "Languages with a finite antidictionary: growth-preserving transformations and available orders of growth"
- 2012 37th International Symposium on Mathematical Foundations of Computer Science (MFCS 2012). Bratislava, Slovakia. "On two stronger versions of Dejean's conjecture"
- 2012 14th Mons Days of Theoretical Computer Science (JM 2012). Louvain-la-Nuove, Belgium. "Binary patterns avoided by the binary cube-free words"
- 2010 14th International Conference on Developments in Language Theory (DLT 2010). London Ontario, Canada. "On the existence of minimal  $\beta$ -powers"
- 2010 5th International Computer Science Symposium in Russia (CSR 2010). Kazan, Russia. "Growth of power-free languages over large alphabets"
- 2010 10th Northern Combinatorial Conference (NORCOM 2010). Reykjavik, Iceland. "Graphs and words: there and back again"

- 2009 13th International Conference on Developments in Language Theory (DLT 2009). Stuttgart, Germany. "Two-sided bounds for the growth rates of power-free languages"
- 2008 12th Mons Days of Theoretical Computer Science, Mons, Belgium. "On the growth rate of complexity of threshold languages"
- 2008 3rd International Computer Science Symposium in Russia (CSR 2008). Moscow, Russia. "Combinatorial complexity of regular languages"
- 2006 10th International Conference on Developments in Language Theory (DLT 2006), Santa-Barbara, USA. "Factorial Languages of low combinatorial complexity"
- 2004 Discrete Analysis and Operation Research (DAOR'04), Novosibirsk, Russia. "Combinatorial complexity of rational languages"
- 2003 4th International Conference on Combinatorics on Words (WORDS'03), Turku, Finland. "The interaction property of periods for partial words"
- 2001 26th International Symposium on Mathematical Foundations of Computer Science (MFCS 2001), Marianske Lasne, Chezh Republic. "On the periods of partial words"
- 2000 International Colloquium on Semigroups, Szeged, Hungary. "Syntactic semigroups of languages closed under factors"
- 1999 2nd International Conference on Semigroups with Applications, St-Petersburg, Russia. "Binary Z-words of a given local exponent"
- 1997 International Algebraic Conference dedicated to the centennial of the birthday of D.K. Faddeev, St.-Petersburg, Russia. "On infinite binary words"
- 1995 International Conference on Semigroups with Applications, including Semigroup Rings, in honour of E.S. Lyapin, St-Petersburg, Russia. "Galois connection in avoidability theory'

## Active Service to the Academic Community

## **Organizing Events**

- 2014 18th International Conference on Developments in Language Theory (DLT 2014), OC member
- 2014 7th School "Computer Science Days in Ekaterinburg", OC chair
- 2013 8th International Computer Science Symposium in Russia (CSR 2013), chair
- 2013 6th School "Computer Science Days in Ekaterinburg", OC member

## Program Committee Membership

- 2017 28th Annual Symposium on Combinatorial Pattern Matching (CPM 2017), PC member
- 2017 19th International Conference on Descriptional Complexity of Formal Systems (DCFS 2017), PC member
- 2015 40th International Symposium on Mathematical Foundations of Computer Science (MFCS 2015), PC member
- 2014 18th International Conference on Developments in Language Theory (DLT 2014), PC co-chair
- 2014 7th School "Computer Science Days in Ekaterinburg", PC chair
- 2013 9th International Conference on Words (WORDS 2013), PC member
- 2013 6th School "Computer Science Days in Ekaterinburg", PC member
- 2012 7th International Computer Science Symposium in Russia (CSR 2012), PC member
- 2004 Workshop on Words Avoidability, Complexity, and Morphisms (WACAM 2004), PC member

## Editorial Board Membership

2005–2011: Editorial board secretary for Proceedings of Ural State University, Ser. Computer Science and Information Technology

2015– : Editorial board member for Ural Mathematical Journal

## Paper Reviewing

Article Reviewer for: Discrete Applied Mathematics, Discrete Mathematics, Theoretical Computer Science, European Journal of Combinatorics, Electronic Journal of Combinatorics, Theory of Computing Systems, Information and Computation, International Journal of Foundations of Computer Science, Fundamenta Informaticae, RAIRO Theoretical Informatics and Applications, Semigroup Forum, Journal of Integer Sequences, Algebra and Logic, Siberian Mathematical Journal, Information Transmission Problems, Discrete Analysis and Operation Research.

*Conference Paper Reviewer* for: Symposia on Theoretical Aspects of Computer Science (STACS), International Symposia on Mathematical Foundations of Computer Science (MFCS), European Symposia on Algorithms (ESA), International Computer Science Symposia in Russia (CSR), Annual Symposia on Combinatorial Pattern Matching (CPM), International Workshops on Combinatorial Algorithms (IWOCA), International Conferences on Developments in Language Theory (DLT), International Conferences on Words (WORDS).

Reviewer for Mathematical Reviews (MSC 68Q70, 68R15).

## Student Scientific Contests Management

- 2016 Alan Turing Contest in Theoretical Computer Science and Discrete Mathematics for Russian-speaking students (TuCo 2016), Jury member
- 2016 All-Russian Möbius Contest in Mathematics for students, Reveiwer

## Teaching and Supervision

## Courses

I teach math and theoretical CS mainly to the students, specialising in computer science, computer security, and pure mathematics in the Institute of Mathematics and Computer Science of the Ural Federal University. I also have some experience in giving math lectures to students studying physics and even philosophy. Below is the list of courses I read on regular basis during the last several years. The last five courses from this list were designed and invented in the Ural Federal University by me.

Algebra and Geometry I, II. A course for 1st year students in CS and CSec. Covers basics on sets and binary relations, analytic geometry, complex numbers, polynomials, matrices and linear algebra (including advanced topics).

*Discrete Mathematics.* A course for 2nd year students in CS and CSec. Covers combinatorics, graphs, boolean functions, and some questions of set theory and general algebra.

Automata Theory. A course for 2nd year students in CS and CSec. Covers a standard piece of theory of finite automata and regular languages and some advanced topics, like finite transducers.

*Grammars and Compiler Theory.* A course for 3rd year students in CS and CSec. Covers the theory of context-free languages and its use for the front end of compilers (LR-parsers, syntax-directed definitions, etc).

*Theory of Algorithms.* A course for 3rd year students in CS and CSec. Covers standard topics on decidability and complexity.

*Combinatorics of Words.* A course of choice for 3rd/4th year students in Math and CS. Covers topics related to periodicity, (un)avoidable repetitions and complexity of words (strings).

String Algorithms. A course of choice for 3rd/4th year students in Math and CS. Covers pattern matching problems and indexing data structures.

*Design and Analysis of Algorithms.* A course for grad students in CS. Covers precise, approximate, and randomized efficient algorithms, together with some complexity issues.

Lossless Data Compression. A course for grad students in CS. Covers basics of information theory, main coding techniques for texts and numbers, and practical compression methods (statistical, dictionary-based, BWT-based).

*Nonnegative Matrices and Applications.* A course of choice for grad students in Math and CS. Covers the Perron-Frobenius theory and some of its applications to graphs, regular languages, and Markov processes.

Apart from the above list, I read courses in General Algebra (for Math students), Differential Geometry (for Math students), Cryptography I (for CSec students), Algebra and Geometry (for Physics students), Higher Mathematics (for Philosophy students).

## Graduate Supervision (Ural State/Federal University)

## Ph.D. Students (Completed)

No.	Name	Period	Title of Thesis
6	Dmitry Kosolobov	2013-2016	Efficient algorithms for the study of regularities in strings
5	Mikhail Rubinchik	2013-2016	Computational complexity of some string processing problems
4	Elena Petrova	2012 - 2016	On combinatorial properties of power-free languages
3	Irina Gorbunova	2009-2013	Constructing and enumerating extremal power-free words
2	Andrei Plyushchenko	2008-2011	Combinatorial properties of Burnside semigroups
1	Yulia Gamzova	2002-2006	Combinatorial properties of partial words

## Master Students (Completed)

No.	Name	Period	Research Area
26	Maria Kiskachi	2014-2016	String algorithms
25	Anton Fedorov	2014-2016	Graph theory
24	Oleg Merkurev	2014-2016	String algorithms
23	Vladimir Plyashkun	2013-2015	Compiler theory
22	Tatiana Sharypova	2013-2015	Combinatorics of words
21	Vitaly Ishkulov	2013-2015	Combinatorics of words
20	Mikhail Rubinchik	2011 - 2013	String algorithms
19	Igor Tunev	2011 - 2013	Combinatorics of words
18	Alexander Golubev	2011 - 2013	Natural language processing
17	Lidia Idiatulina	2010-2012	Graph theory
16	Elena Petrova	2010-2012	Combinatorics of words
15	Alexei Samsonov	2009-2011	String algorithms
14	Alina Voskova	2009-2011	Graph theory
13	Konstantin Likhomanov	2008-2010	Combinatorics of words
12	Nikita Vostretsov	2007 - 2009	Natural language processing
11	Dmitry Ivankov	2007 - 2009	Compiler theory
10	Irina Gorbunova	2007 - 2009	Combinatorics of words
9	Anastasiya Tkach	2007 - 2009	Automata theory
8	Vyacheslav Opletin	2006-2008	Data compression
7	Andrei Plyushchenko	2006-2008	Combinatorics of words
6	Dmitry Didenko	2006-2008	Automata theory
5	Tatiana Senchonok	2004-2006	Combinatorics of words
4	Alexandr Lukanin	2003-2005	Data compression
3	Olga Toporkova	2003-2005	Pseudorandom sequences
2	Elena Nikolaeva	2001 - 2003	Data compression
1	Yulia Konovalova	2000-2002	Combinatorics of words

## Ph.D. Students (Current)

No.	Name	Period	Research Area
1	Oleg Merkurev	2016-	String algorithms

## Master Students (Current)

No.	Name	Period	Research Area
6	Olga Karpova	2016-2018	Combinatorics of words
5	Daniil Gasnikov	2016-2018	Combinatorics of words
4	Anton Chaplygin	2016-2018	String algorithms
3	Konstantin Startsev	2016-2018	Combinatorics of words
2	Aleksandr Chukharev	2016-2018	Combinatorics of words
1	Aleksandr Dubikovskii	2016-2018	String algorithms

# Research/Project Funding Received

(For convenience, all amounts are converted into USD)

Type	Awarded	Period	Project Title	Funding	Annual	Total	Involvement
Research	2016	2016-	Algorithmics and combi-	Russian Founda-	\$12000	\$36000	Co-
Grant		2018	natorics of strings and fi-	tion of Basic Re-			Investigator
			nite automata	search			
Teaching	2014	2015	Combinatorics of words	Dynasty Founda-	\$4000	\$4000	Principal
Grant			with applications	tion and Yandex			Investigator
Research	2014	2014-	Center of excellence in	Ministry of Sci-	\$120000	\$360000	Co-
Project		2016	theoretical computer sci-	ence and Educa-			Investigator
			ence	tion of Russia			
Research	2013	2013-	New challenges in formal	Russian Founda-	\$16000	\$48000	Co-
Grant		2015	languages and automata	tion of Basic Re-			Investigator
			theory	search			
Fellow-	2011	2011	Ebco Eppich Visitor Fel-	Simon Fraser	\$6000	\$6000	Principal
ship			lowship	University			Investigator
Research	2010	2010-	Dynamics of finite au-	Russian Founda-	\$14000	\$42000	Co-
Grant		2012	tomata and regular lan-	tion of Basic Re-			Investigator
			guages	search			
Research	2009	2009-	Algebraic systems and	Ministry of Sci-	\$80000	\$240000	Co-
Project		2011	their application to	ence and Educa-			Investigator
			computer science and	tion of Russia			
			bioinformatics				
Research	2010	2010-	Basic problems in the the-	Russian Founda-	\$15000	\$45000	Co-
Grant		2012	ory of algebraic systems	tion of Basic Re-			Investigator
			and their application to	search			
			computer science				
Research	2006	2006-	Combinatorics of words	Russian Founda-	\$12000	\$24000	Co-
Grant		2007	and automata and its ap-	tion of Basic Re-			Investigator
			plications to computer sci-	search			
			ence and bioinformatics				
Research	2006	2006	Combinatorial characteri-	Science and Inno-	\$14000	\$14000	Principal
Grant			zation of formal languages	vation Agency of			Investigator
				Russia			

Research	2005	2005	Periodicity, avoidability,	Science and Inno-	\$12000	\$12000	Principal
Grant			and complexity of formal	vation Agency of			Investigator
			languages: research stay	Russia			
			at the University of Turku				
Scholar-	1994	1995 -	Soros scholarship for post-	Soros "Open So-	\$1800	\$5400	Principal
ship		1997	graduate students	ciety" Program			Investigator