Outstanding Challenges in Combinatorics on Words (12w5068)

James Currie (University of Winnipeg),
Jeffery Shallit (University of Waterloo)

Feb. 19 – Feb. 24, 2012

1 Overview of the Field

Combinatorics on words is a relatively new area of research in discrete mathematics. It studies the properties of words (sequences), either finite or infinite, over a finite alphabet. The perspective on words can be variously algebraic, combinatorial, or algorithmic. The field is characterized by its manifold connections to topics, not only in mathematics, but also in other scientific disciplines. Inside mathematics, connections exist to certain parts of algebra (e.g., combinatorial group theory and semigroups), probability theory, number theory, and discrete symbolic dynamics. In other sciences, connections exist to crystallography, and to DNA sequencing. Combinatorics on words has also been notably connected to and motivated by theoretical computer science, e.g., automata theory and pattern matching algorithms.

The Norwegian number theorist A. Thue [22, 23] was the first to study combinatorial problems on sequences of symbols for their own sake. Among other things, he established the existence of an infinite word on two symbols (three symbols) having no cubes (squares, respectively) as factors. Here, a cube refers to three consecutive repetitions of the same word. Thue's research was the origin of research on pattern avoidability in words. Such results have had profound applications, e.g., in Burnside-type questions in algebra.

Later, over several decades, several papers on combinatorics on words appeared, mainly as tools of solving problems on different areas. A typical example is a paper by M. Morse [17] in the 1930's, where the motivation arose from symbolic dynamics. This work initiated research on the subword complexity of words. Another example is provided by the Russian mathematician A. I. Shirshov who while studying polynomial rings, discovered a fundamental unavoidability result of words, subsequently known as Shirshov's theorem [21].

Combinatorics on words is now listed in the AMS Mathematics Subject Classification as a distinct topic, under the code 68R15. The emergence of fundamental results on sequences in different domains of mathematics is one of the motivations for this identification of combinatorics on words as an area of research. Van der Waerden's theorem [24] and the periodicity lemma of Fine and Wilf [9] are further examples of this phenomenon. The former was stated in terms of colorings of the natural numbers, and the latter as a periodicity criterion for real functions. However, their most natural formulation is as properties of words. Systematic research on words started in the 1950's in two different research communities: the Moscow and the French schools. The Moscow school was initiated by P. S. Novikov and S. I. Adjan in connection to their research on the Burnside problem and the establishment of the undecidability of the word problem for groups [1]. At around the same time M. P. Schützenberger created the French school as part of his research on the theory of codes [20]. Both schools can be viewed as pioneers of the field. The decidability of the satisfiability problem for word equations by G. S. Makanin [16] is an important sample outcome. Over the last 20 years or so combinatorics on words has developed into a quickly growing topic of its own; a few textbooks [13, 14, 15, 2] have also emerged as very influential.

2 Recent Developments

Several striking results have been achieved in this area recently. Among these are the resolution of long-standing problems and conjectures:

- A 1972 conjecture by F. Dejean [8] stated a precise bound on the size of unavoidable repetitions in infinite words. This conjecture was finally confirmed through the work of M. Rao, J. Currie, N. Rampersad, and A. Carpi [19, 7, 4].
- The centralizer of a language is the maximal language commuting with it. The question, raised in 1970 by J. H. Conway [6], whether the centralizer of a rational language is always rational has been negatively answered with a celebrated result [12]. In fact, even complete co-recursively enumerable centralizers exist for finite languages.
- The satisfiability of word equations with constants is in PSPACE [18]. It follows from the proof of that result that the satisfiability of word equations with constants is in NP if one shows that the minimal solutions of a word equations are single exponential in the size of the equation if they exist.
- The solution of J.-P. Duval's 1982 conjecture and the Ehrenfeucht-Silberger problem about the relation between the period of a word and the maximum length of its unbordered factors by T. Harju, S. Holub, and D. Nowotka in 2004 settled long-standing questions [10, 11].
- Does there exist an infinite word over a finite subset of N such that no three consecutive blocks of the same size and the same sum exist? G. Pirillo and S. Varricchio raised that question 1994 in the context of semigroup theory. L. Halbeisen and N. Hungerbuhler formulated that problem in different terminology in 2000 independently of G. Pirillo and S. Varricchio. Just recently that question was affirmatively answered by J. Cassaigne, J. D. Currie, L. Schaeffer, and J. Shallit [5].

In addition, tools in several subareas are clearly coming to maturity. At one point, the connections between combinatorics on words and transcendence results seemed to be one-way only, and somewhat ad hoc; today, this connection is better understood, and several tools have emerged in this intersection of discrete mathematics with algebra. Major progress has also been made on variations of the run-length problem, which ties together combinatorics on words with ideas from data compression. As a final example, properties of automatic sequences, expressed in a certain logic, have extremely recently been shown to be decidable.

3 Presentation Highlights

3.1 Run-length and maximal exponent problems

3.1.1 Videotaped lecture

For the first of our videotaped lectures, *Maximal Exponent Repeats*, **Maxime Crochemore** presented an overview of the run-length problem and related questions. This talk was a summary of basic issues related to repetitions in strings, concentrating on algorithmic and combinatorial aspects. This area is important both from theoretical and practical point of view. Repetitions are highly periodic factors (substrings) in strings and are related to periodicities, regularities, and compression. The repetitive structure of strings leads to higher compression rates, and conversely, some compression techniques are at the core of fast algorithms for detecting repetitions. There are several types of repetitions in strings: squares, cubes, and maximal repetitions also called runs. For these repetitions, we distinguish between the factors (sometimes qualified as distinct) and their occurrences (also called positioned factors). The combinatorics of repetitions is a very intricate area, full of open problems. For example we know that the number of (distinct) primitively-rooted squares in a string of length n is no more than $2n\theta(\log n)$, and is conjectured to be n, and that their number of occurrences can be $\theta(n \log n)$. Similarly we know that there are at most $1.029n$ and at least $0.944n$ maximal repetitions and the conjecture is again that the exact bound is n.

3.1.2 Some other notable talks related this area

Golnaz Badkobeh gave a talk *Fewest repetitions vs maximal-exponent powers in infinite binary words*
Abstract: The exponent of a word is the ratio of its length over its smallest period. The repetitive threshold $r(a)$ of an a-letter alphabet is the smallest rational number for which there exists an infinite word whose finite factors have exponent at most $r(a)$. This notion was introduced in 1972 by Dejean who gave the exact values of $r(a)$ for every alphabet size a as it has been eventually proved in 2009. The finite-repetition threshold for an a-letter alphabet refines the above notion. It is the smallest rational number FR$t(a)$ for which there exists an infinite word whose finite factors have exponent at most FR$t(a)$ and that contain a finite number of factors with exponent $r(a)$.

It is known from Shallit (2008) that FR$t(2) = 7/3$. With each finite-repetition threshold is associated the smallest number of $r(a)$-exponent factors that can be found in the corresponding infinite word.

It has been proved by Badkobeh and Crochemore (2010) that this number is 12 for infinite binary words whose maximal exponent is $7/3$.

In this article we give some new results on the trade-off between the number of squares and the number of maximal-exponent powers in infinite binary words, in the three cases where the maximal exponent is $7/3$, $5/2$, and 3. These are the only threshold values related to the question.

Jamie Simpson spoke on *Matching fractions*
Abstract: If x is an infinite word and k is a non-negative integer then the kth matching fraction of x, written $\mu(k)$, is the following limit, if it exists:

$$\mu(k) = lim_{n\to\infty} \frac{|\{i \in [1, n] : x[i] = x[i + k]\}|}{n}. \tag{1}$$

The idea of a matching fraction was suggested by Peter Pleasants and has been investigated by Keith Tognetti and colleagues. In particular they showed that if x is a Sturmian word on alphabet $\{a, b\}$ whose density of as is α then

$$\mu(k) = \max(|1 - 2\{\alpha\}|, |1 - 2\{k\alpha\}|). \tag{2}$$

Here $\{\alpha\}$ is the fractional part of α.

I will review what's known about matching fractions and suggest some areas for investigation.

3.2 Diophantine approximation

Monday morning was devoted to an exposition of topics relating combinatorics on words and automatic sequences to diophantine approximation.

Boris Adamczewski gave a lecture *Combinatorics on words and Diophantine approximation*
Abstract: A very fruitful interplay between combinatorics on words and Diophantine approximation comes up with the use of numeration systems.

Finite and infinite words occur naturally in Number Theory when one considers the expansion of a real number in an integer base or its continued fraction expansion. Conversely, with an infinite word w on the finite alphabet $\{0, 1, \ldots, b - 1\}$ one can associate the real number ξ_w whose base-b expansion is given by w. Many problems are then concerned with the following question: how the combinatorial properties of the word w and the Diophantine properties of the number ξ_w may be related?

In this talk, I will survey some of the recent advances on this topic. I will also try to point out new challenges.

A closely related talk was given by **Jason Bell** speaking *On automatic sequences*.

3.3 Logic and decision problems

Several workshop members work in the area of decision problems and logic, and the existence or non-existence of algorithms for solving problems in combinatorics on words. Of course, word problems go back to the beginning of the theory of algorithms, since the Post Correspondence Problem (PCP) is phrased in the laguage of strings, or alternatively, as a question related to morphisms. Some talks addressed issues of this flavour:

Jeffrey Shallit spoke on *On k-automatic sets of rational numbers*
Abstract: In this talk I will describe how automata can accept sets of rational numbers. Applications include deciding various questions about automatic sequences and the enumeration of various properties. There are many open questions.

Yuri Matiyasevich, famous for his solution of Hilbert's tenth problem, gave a brief lecture on *Arithmetization of words and related problems*

3.4 Complexity measures

As noted in our overview of the field, the work of Morse and Hedlund led to the study of factor complexities of various classes of words. This remains a major theme in combinatorics on words, with several open problems of long standing.

3.4.1 Videotaped lecture

A presentation by **Amy Glen** entitled *On a generalisation of trapezoidal words* was recorded:
Abstract: The *factor complexity function* $C_w(n)$ of a finite or infinite word w counts the number of distinct *factors* of w of length n for each $n \geq 0$. A finite word w of length $|w|$ is said to be *trapezoidal* if the graph of its factor complexity $C_w(n)$ as a function of n (for $0 \leq n \leq |w|$) is that of a regular trapezoid (or possibly an isosceles triangle); that is, $C_w(n)$ increases by 1 with each n on some interval of length r, then $C_w(n)$ is constant on some interval of length s, and finally $C_w(n)$ decreases by 1 with each n on an interval of the same length r. Necessarily $C_w(1) = 2$ (since there is one factor of length 0, namely the *empty word*), so any trapezoidal word is on a binary alphabet. Trapezoidal words were first introduced by A. de Luca (1999) when studying the behaviour of the factor complexity of *finite Sturmian words*, i.e., factors of infinite "cutting sequences", obtained by coding the sequence of cuts in an integer lattice over the positive quadrant of $\mathbf{R}^2$ made by a line of irrational slope. Every finite Sturmian word is trapezoidal, but not conversely. However, both families of words (trapezoidal and Sturmian) are special classes of so-called *rich words* – a wider class of finite and infinite words characterised by containing the maximal number of palindromes – recently introduced by the speaker, together with J. Justin, S. Widmer, and L.Q. Zamboni (2009).

In this talk, I will introduce a natural generalisation of trapezoidal words over an arbitrary finite alphabet A consisting of at least two distinct letters, called *generalised trapezoidal words* (or *GT-words* for short). In particular, I will discuss some combinatorial properties of this new class of words when $|A| \geq 3$ and I will show that, unlike in the binary case ($|A| = 2$), not all GT-words are rich in palindromes, but we do have a neat characterisation of those that are.

This is joint work with Florence Levé (Université de Picardie – Jules Verne, France).

3.4.2 Some other notable talks related this area

Words form basic objects in a variety of areas. Free groups, for example, are naturally viewed as sets of words, related by combinatorial rules. Within combinatorics on words, several now classical constructions comprise standard examples, notably the Thue-Morse word and the Sturmian words. However, open problems concerning these classes remain open. The following trio of talks by young researchers addressed various classical objects, with a unifying theme of complexity (of factors, or otherwise):

Thomas Stoll (Université d'Aix-Marseille) addressed *Thue-Morse at polynomial subsequences*
Abstract: Consider the Thue-Morse sequence (t_n) on the symbols $\{0, 1\}$. A celebrated result of Gelfond

(1967/68) gives an asymptotic formula for the number of 0's and 1's in any fixed linear subsequence of TM, i.e., (t_{an+b}). In quite recent work, Mauduit and Rivat (2009) found precise formulas for quadratic polynomials. The cases of cubes and higher-degree polynomials remain elusive so far. From work of Dartyge and Tenenbaum (2006) it follows that there are $\gg N^{2/h!}$ symbols "0" (or "1") in any subsequence indexed by a polynomial of degree h. The aim of the present talk is to give an overview about the various results known in this area, pose some (old and new) conjectures, and – to improve by elementary/combinatorial means the general lower bound to $\gg N^{4/(3h+1)}$.

Eric Rowland spoke on *Counting equivalence classes of words in F_2*
Abstract: In the last decade several papers have appeared concerning the size of an equivalence class of words in a free group under its automorphism group. A central theme of the area is that information about the equivalence class of a word can be obtained from statistics of its (contiguous) subwords. Here we are interested in the free group on 2 generators. We give a new characterization of words of minimal length, and we introduce a natural operation that "grows" words from smaller words. The growth operation gives rise to a notion of maximally minimal words (in a way that can be made precise), which we call root words. Equivalence classes containing root words have special structure, and the hope is that understanding this structure will lead to an exact enumeration of equivalence classes in F_2 containing a minimal word of length n.

Julien Leroy gave a lecture entitled *The S-adic conjecture)*
Abstract: An infinite word is S-adic if it can be obtained by successive iterations of morphisms belonging to the set S. Sturmian words are well-known examples of S-adic words with card(S) $= 4$. The S-adic conjecture tries to determine the link that should exist between S-adicity and sub-linear factor complexity. More precisely, it says that there is a stronger notion of S-adicity which is equivalent to sub-linear factor complexity, i.e., an infinite word would have a sub-linear complexity if and only if it is "strongly S-adic".

In this talk, I will present some recent results about that conjecture. First I will present some examples that allow to reject some natural ideas that one could have. Then I will briefly explain a general method to compute an S-adic expansion of any uniformly recurrent infinite word with sub-linear complexity. That method allows to solve the conjecture in the particular case of uniformly recurrent infinite words with first difference of complexity bounded by 2.

3.5 Pattern avoidance

A classical topic in combinatorics on words is the avoidance of patterns such as repetitions (squares, cubes, overlaps) and images of words. Work of this sort contributed to the solution of the Burnside problem for groups. However, analogously to number theory, the study of pattern avoidance is a source of simply stated, yet apparently intractable problems. While repetitions involve the concatenation of identical factors, several researchers are now interested in generalizations where subsequent factors are equivalent up to order, or under various transformations. Several stimulating talsk on this theme were presented:

Juhani Karhumäki presented a lecture *Combinatorics on words and k-abelian equivalence*
Abstract: We define a new equivalence relation on words which is properly in between the equality and commutative (abelian) equality. We call it k-abelian equality. We say that two words are k-abelian equal if they contain each factor of length k equally many times, and in addition start with a common prefix of length k-1. This allows to have better and better approximations of problems based on equality of words. We report basic properties of these equivalence relations, and analyze k-abelian variants of some well known problems on words. In particular we consider problems dealing local and global regularities in infinite words, as well as the existence of certain type of repetition-free words. It turns out that there exist a lot of challenging open problems in this area.

Antonio Restivo gave a talk showing the state of the art in an area related to the critical factorization theorem of Fine and Wilf in his lecture *Sturmian Words and Critical Factorization Theorem*
Abstract: We prove that characteristic Sturmian words are extremal for the Critical Factorization Theorem (CFT) in the following sense. If $px(n)$ denotes the local period of an infinite word x at point n, we prove that x is a characteristic Sturmian word if and only if $px(n)$ is smaller than or equal to $n + 1$ for all $n \geq 1$ and it is equal to $n + 1$ for infinitely many integers n.

This result is extremal with respect to the CFT since a consequence of the CFT is that, for any infinite recurrent word x, either the function $px(n)$ is bounded, and in such a case x is periodic, or $px(n) \geq n+1$ for infinitely many integers n. As a byproduct of the techniques used in the paper we extend a result of of Harju and Nowotka stating that any finite Fibonacci word f_n for $n \geq 5$ has only one critical point. Indeed we determine the exact number of critical points in any finite standard Sturmian word.

Young researcher **Robert Mercas** presented a different perspective on results of Fine and Wilf: *On Pseudo-repetitions in words*
Abstract: The notion of repetition of factors in words was studied already from the beginnings of the combinatorics on words area. One of the recent generalizations regarding this concept was introduced by L. Kari et al., and considers a word to be an f-repetition if it is the iterated concatenation of one of its prefixes and the image of this prefix through an anti-/morphic involution f. In this paper, we extend the notion of f-repetitions to arbitrary anti-/morphisms, and investigate a series of algorithmic problems arising in this context. Further, we present a series of results in the fashion of the Fine and Wilf theorem for f-repetitions, when f is an iso(anti)morphism.

Several other speakers contributed to this topic:

Dirk Nowotka spoke on *Avoidability under Permutations*; **Arturo Carpi** (Universita Degli Studi Di Perugia) gave a lecture on *Unrepetitive walks in digraphs (and the repetition threshold)*; **Sébastien Ferenczi** presented *Word combinatorics of interval-exchange transformations for every permutation***Anna Frid** gave a brief problem involving *Morphisms on Permutations*

3.6 Word equations

Several speakers addressed the state of decision procedures for word equations.

Štěpán Holub spoke on *Length types of word equations*
Abstract: One of the 'rules of thumb' for solving a word equation is to guess the length vector of the solution and then construct the corresponding equivalence on the positions. The talk will comment on this method and indicate some less trivial implications.

Luca Zamboni presented work tying combinatorics on words to ultrafilter constructions in algebra: *Partition regularity and words*
Abstract: Van der Waerden's theorem states that given a finite partition of the natural numbers N, at least one element of the partition contains arbitrarily long arithmetic progressions (i.e., is a.p.-rich). In fact, in the hypothesis N can be replaced by any subset X of N which is itself a.p.-rich. In other words, the property of being a.p.-rich is partition regular, i.e., cannot be destroyed under finite partition. However given a partition regularity property P (e.g., being a.p.-rich) and a subset X of N (e.g., $X =$set of all prime numbers), the question of determining whether X has property P is often quite difficult (even if either X or its complement must have property P). In the context of combinatorics on words it is natural to consider X to be the set of all occurrences of a given factor in some nice infinite word (i.e., the set of all occurrences of 0 in the Thue-Morse word, or the set of all occurrences of 0100 in the Fibonacci word). In this talk we will consider different partition regularity properties in the context of words. It turns out that many well known words (including Sturmian words and words generated by substitution rules) provide a rich setting for studying different partition regularity properties. And conversely, some deep dynamical and or arithmetical questions concerning these words may be reformulated in terms of partition regularity. One such example is the strong coincidence condition (SCC) conjectured for irreducible primitive Pisot substitutions. We will describe an equivalent reformulation of SCC whose difficulty lies in the arithmetic properties of the associated Thomas-Dumont numeration systems. This is based on joint work with M. Bucci and S. Puzynina.

Volker Diekert , in his talk *Local Divisors*, highlighted the relevance of combinatorics on words to group-theoretic questions.
Abstract: In my talk I will speak about several new results and simpler proofs of known results in formal language theory using the concept of a local divisor. The results encompass recent joint work with Manfred

Kufleitner and Benjamin Steinberg on Krohn-Rhodes Theorem.

3.7 Other talks

Researchers at the workshop were extremely generous in volunteering talks. We also heard presentations by

- **Nicolas Bédaride** *Piecewise isometries and words*
- **Srecko Brlek** *The last talk on the Kolakoski sequence*
- **Svetlana Puzynina** *Locally catenative sequences and Turtle graphics*

4 Open problems

A session on open problems identified several questions:

1. If u, v, w are primitive words, define $p(u, v, w)$ to be the integer k such that
$$(u^* \text{ shuffle } v^*) \cap w^* = (w^k)^*.$$
Then define $P(u, v) = \{p(u, v, w) : w \text{ primitive }\}$. Characterize the u, v such that $P(u, v) = \{0, 1\}$.
2. If x is a prefix of infinitely many square-free words over $\{1, 2, 3\}$ and y is a suffix of infinitely many square-free words over $\{1, 2, 3\}$, must there exist some squarefree word xuy over $\{1, 2, 3\}$?
3. Does the paperfolding word contain arbitrarily large Abelian powers?
4. Can the iterated hairpin completion of a singleton w (a) be regular (b) be context-free but not regular?
5. Improve the bounds on run lengths and sums of exponents of runs:
$$\begin{array}{rcl} 0.9445757|x| \leq & \text{runs}(x) & \leq 1.029|x| \\ 2.035|x| \leq & \text{sum of exponents of runs(x)} & \leq 4.1|x| \\ 0.406|x| \leq & \text{3-runs}(x) & \leq 0.5|x|. \end{array}$$
6. Are 2-Abelian cubes 2-avoidable?

5 Scientific Progress Made and Outcome of the Meeting

Several researchers have reported papers and research progress related to this one week workshop. In particular:

- The paper S. Ferenczi *The self-dual induction for every interval exchange transformation* (in preparation) was improved through several discussions in Banff.
- The paper S.Ferenczi, L. Zamboni *Clustering words* (http://arxiv.org/abs/1204.1541) was started during the (homeward) train trip from Banff and Montreal.
- Steffen Kopecki and Volker Diekert worked at BIRS on the combinatorics of hairpin completions. These combinatorial problems about words are inspired by biochemical processes in DNA computing where hairpin formations arise naturally. During the workshop they sharpened some of their results which became part of the revised version of their paper
  - Volker Diekert, Steffen Kopecki, Victor Mitrana *Deciding Regularity of Hairpin Completions of Regular Languages in Polynomial Time*.

The paper was just accepted for publication in *Information and Computation* on April 17, 2012 and will contain an acknowledgement to the inspiring workshop at BIRS.

- The recent BIRS workshop was a great opportunity for Boris Adamczewski to advance his collaboration with Jason Bell. In particular, it allowed them to
  - (Almost) Finish the writing of a joint paper on diagonals of multivariate algebraic functions.
  - Discuss a current project on Mahler's functions
  - Discuss various new questions. They also restarted a joint project with Berthé and Zamboni.
- In addition Bell was able to
  - Work on characterization of subsets of $\mathbb{R}$ accepted by Buchi automata with respect to two multiplicatively independent bases (with Julien Leroy and Emilie Charlier.
  - Begin work on various questions of Jeff Shallit on k-automatic sequences with specific properties.
- Štěpán Holub solved the workshop open problem on the paperfolding word, shortly after his return from BIRS.
- Thomas Stoll also participated in discussion regarding the paperfolding word, and also
  - Answered a question of Shallit asking whether it is possible to give a bound for $\min(n : t_{k_1 n} = e_1, t_{k_2 n} = e_2)$ where t_n is the Thue-Morse sequence and k_1, k_2 are arbitrary distinct integers and e_1, e_2 all of the four possibilities. He got the result during the BIRS workshop and they are now in progress toward finding the natural generalized result (with two students, one in Marseille and one from Stanford).
  - Also with Shallit, Stoll found during BIRS the proof of a conjecture of Eric Rowland that the 2-kernel oft_{n+l} is of size $f(l)$ where $f(l)$ satisfies some nice explicit recursive relations and is k-regular.
  - Researchers from France (Stoll, Cassaigne, Ochem) were also inspired to think about the (difficult) problem of avoiding sumsquares in words over finite alphabets.

Researchers were extremely enthusiastic about the workshop, and the wonderful atmosphere at BIRS. Thanks to BIRS for putting this resource at the disposal of the international mathematical community.

References

[1] S. Adjan & P.S. Novikov, Infinite periodic groups I–III. (Russian) *I*zv. Akad. Nauk SSSR Ser. Mat. **3**2 (1968).

[2] Jean-Paul Allouche and Jeffrey Shallit, *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, 2003.

[3] Jean Berstel, *A*xel Thue's papers on repetitions in words: a translation, Publications du LCIM **2**0, Université du Quebec à Montréal (1994)

[4] A. Carpi, On Dejean's conjecture over large alphabets, Theoret. Comput. Sci. 385 (2007) 137–151.

[5] J. Cassaigne, J. Currie, L. Schaeffer, J. Shallit, Avoiding Three Consecutive Blocks of the Same Size and Same Sum, arXiv:1106.5204v3

[6] *R*egular Algebra and Finite Machines, Chapman Hall, 1971.

[7] J. D. Currie, N. Rampersad, *A* proof of Dejean's conjecture, Math. Comp. **8**0 (2011), 1063–1070.

[8] F. Dejean, Sur un théorème de Thue, J. Combin. Theory Ser. A 13 (1972) 90–99.

[9] N. J. Fine and H. S. Wilf, Uniqueness theorem for periodic functions. *P*roc. Amer. Math. Soc., /textbf 16 (1965) 109–114.

[10] T. Harju and D. Nowotka, Periodicity and unbordered words: A proof of the extended Duval conjecture, *J*. ACM **5**4(4) (2007).

[11] S. Holub, A proof of the extended Duval conjecture, *T*heor. Comput. Sci., **3**39 (2005) 61–67.

[12] M. Kunc, Regular solutions of language inequalities and well quasi-orders, *P*roc. of ICALP 2004, 870–888, 2004.

[13] M. Lothaire, *C*ombinatorics on Words, Encyclopedia of Mathematics and its Applications **1**7, Addison-Wesley, Reading (1983).

[14] M. Lothaire, *A*lgebraic Combinatorics on Words, Cambridge University Press, Cambridge (2002).

[15] M. Lothaire, *A*pplied Combinatorics on Words, Cambridge University Press, Cambridge (2005).

[16] G. S. Makanin, The problem of solvability of equations in a free semigroup, *M*at. Sb. (N.S.) **1**03(145) (1977) 147–236

[17] M. Morse, Recurrent geodesics on a surface of negative curvature, *T*rans. AMS **2**2 (1921), 84–100.

[18] W. Plandowski, Satisfiability of word equations with constants is in PSPACE, *J*. ACM **5**1(3) 483–496 (2004).

[19] M. Rao, Last Cases of Dejean's Conjecture, *T*heor. Comput. Sci. **4**12 (2011), 3010–3018.

[20] M. P. Schtzenberger, Une thorie algbrique du codage, *S*éminaire Dubreil-Pisot, 1955–1956, Exposé No. 15.

[21] A. I. Shirshov, *B*ases of free Lie algebras, Algebra i Logika **1**, 1962, 14–19.

[22] Axel Thue, Über unendliche Zeichenreihen, *N*orske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiana No. 7 (1906).

[23] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiana 1 (1912) 1–67.

[24] B. L. van der Waerden, Beweis einer Baudetschen Vermutung, *N*ieuw. Arch. Wisk. **1**5 (1927) 212–216.