Finite-Repetition threshold for infinite ternary words

Golnaz Badkobeh[1], Maxime Crochemore[1,2]

[1]Algorithm Design Group
Department of Computer Science, King's College London
[2]Institut Gaspard-Monge,
Université de Marne-la-Vallée, France

Sep 15, 2011

Overview

Overview

- Fraenkel and Simpson result,
- Dejean's Conjecture on the Repetitive threshold,
- Overlap-free words and number of squares they contain,
- Finite-Repetition Threshold, Shallit's Theorem,
- Finite-Repetition Threshold, main result
- Conclusion
- Generalizing the result for larger alphabet sizes?

Infinite Binary with 3 squares

Repetitions

Example: $\underbrace{abaab}abaabab$ Length=12

Period=abaab

Exponent=$\frac{length}{periodlength} = \frac{12}{5} = 2.4$

Square:exponent even integers. e.g abab

Cube:exponent integers multiple of 3. e.g aaa

Fraenkel and Simpson Theorem

There is an infinite Binary word containing 3 squares 00, 11 and 0101 only.

Example

011000111001011001110001011100101100010111000110 01...

Proof

- Fraenkel, Simpson 1995 (complicated construction)
- Harju, Nowotka 2006

$$\begin{cases} 0 \to 111000110010110001110010 \\ 1 \to 111000101100011100101100010 \\ 2 \to 111000110010110001011100101100 \end{cases}$$

- Ochem 2006 (50-uniform)

$$\begin{cases} 0 \to 00011001011000111001011001110001011100101100010111 \\ 1 \to 00011001011000101110010110011100010110001110010111 \\ 2 \to 00011001011000101110010110001110010111000101100111 \end{cases}$$

- Badkobeh, Crochemore 2010

$0 \to 012, 1 \to 02, 2 \to 1.$

$$\begin{cases} 0 \to 01001110001101 \\ 1 \to 0011 \\ 2 \to 000111 \end{cases}$$

Avoidability on 2 letters

Thue-Morse morphism t

$B = \{0, 1\}$

$$\begin{cases} t(0) & = 01, \\ t(1) & = 10. \end{cases}$$

$t^{\infty}(0)$= $0110100110010110\ldots$.
$t^{\infty}(0)$ is overlap-free.
Overlap is 2^{+}-repetition in the form $auaua$ e.g $\underbrace{abc}\underbrace{abc}a$.

Squares are unavoidable

Avoidability on 3 letters

The morphism f_0 is defined from A to itself

$A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$

$$\begin{cases} f_0(\mathtt{a}) & = \mathtt{abc}, \\ f_0(\mathtt{b}) & = \mathtt{ac}, \\ f_0(\mathtt{c}) & = \mathtt{b}. \end{cases}$$

Iteration and Properties

$f_0(\mathtt{a}) = \mathtt{abc}$,
${f_0}^2(\mathtt{a}) = \mathtt{abcacb}$,

.

.

.

${f_0}^{\infty}(\mathtt{a}) = \mathtt{abcacbabcbacabcacb\ldots}$.

It is known that ${f_0}^{\infty}(\mathtt{a})$ is square-free.

Avoidability on 3 letters

The Dejean's morphism on 3 letters

The Dejean's morphism d is defined from A to itself

$$\begin{cases} d(\mathrm{a}) & = \mathtt{abcacbcabcbacbcacba}, \\ d(\mathrm{b}) & = \mathtt{bcabacabcacbacabacb}, \\ d(\mathrm{c}) & = \mathtt{cabcbabcabacbabcbac}. \end{cases}$$

$d^\infty(\mathrm{a}) = \mathtt{abcacbcabcbacbcacbabcabacabcacbacabacb}\ldots$
The $d^\infty(\mathrm{a})$ is $(\frac{7}{4})^+$-free.

$\frac{7}{4}$ power is unavoidable

The longest word on 3 letters that can avoid $(\frac{7}{4})$ power has length 39.

Repetitive Threshold

- $\mathrm{MaxExp}(x)$= max{ $e \in \mathbb{Q} | u^e$ occurs in x for some u }
- Threshold, t_k for k-letter alphabet is t if
  - Exists infinite string x, $\mathrm{MaxExp}(x) = t$
  - no infinite string y,satisfies $\mathrm{MaxExp}(y) < t$
- Example $t_2 = 2, \ t_3 = \frac{7}{4}$

Dejean's Conjecture on the Repetitive threshold

Dejean's Conjecture

The repetition threshold for a k-letter alphabet (with $k \geq 2$) is:

$$\begin{cases} 2 & \text{for } k = 2 \\ \frac{7}{4} & \text{for } k = 3 \\ \frac{7}{5} & \text{for } k = 4 \\ \frac{k}{k-1} & \textit{otherwise}. \end{cases}$$

Proofs

- *J.J. Pansiot*. A Propos d'une Conjecture de F. Dejean sur les Répétitions dans les Mots.(1984)
- *J. Moulin Ollagnier*. Proof of Dejean's Conjecture for alphabets with $5, 6, 7, 8, 9, 10$ and 11 letters. (1992),
- *A. Carpi*. On Dejean's conjecture over large alphabets.(2007)
- *J. Currie, N. Rampersad*. Dejean's Conjecture holds for $n \geq 27$ (2009)
- *M. Rao*. Last Cases of Dejean's Conjecture. $8 \leq k \leq 34$. (2009).
- *J. Currie, N. Rampersad*. Dejean's Conjecture holds for $n \geq 30$ (2009)

Overlap-free Binary words and number of squares they contain

Observations

Maximum length of an overlap-free binary word with at most s squares.

s	$= 0$	1	2	3	4	5	6	7	8	9	10	11
$\ell(s)$	$= 3$	5	8	12	14	18	24	26	32	38	50	54

s	$= 12$	13	14	15	16	17	18	19	20	21	22	23
$\ell(s)$	$= 66$	78	102	110	134	158	206	222	270	318	414	446

Theorem

[Shallit's theorem (2008)] *For all $t \geq 1$, there are no infinite binary words that simultaneously avoid all squares yy with $|y| \geq t$ and $\frac{7}{3}$ powers.* Consequently there is no infinite overlap-free binary word with a bounded number of squares.

Finite-Repetition Threshold

Finite-Repetition Threshold on k-letter alphabet, $FRt(k)$, is f if

- Exists an infinite k-letter word x with $\mathrm{MaxExp}(x) = f$ and with finitely many t_k-powers.(t_k is Dejean's threshold)
- No infinite k-letter word y with $\mathrm{MaxExp}(y) < f$ contains finitely many t_k-powers.

Example: $k = 2$, $f = \frac{7}{3}$

$FRt(2) = \frac{7}{3}$, Badkobeh and Crochemore (JM2010)

There exists an infinite binary word whose factors have an exponent at most 7/3 and that contains ***12 squares****, the fewest possible.*

12 is the minimum number.

Maximum length of $(7/3)^+$ free binary words with at most s squares.

s	$= 0$	1	2	3	4	5	6	7	8	9	10	11
$\ell(s)$	$= 3$	5	8	12	14	18	24	30	37	43	83	116

Main Result

- The *finite-repetition threshold* of the 3-letter alphabet is its Dejean's repetition threshold, that is, $7/4$.
- The smallest number of $7/4$-powers occurring in a $7/4^+$-power free infinite ternary word is 2, the minimum number.
- The largest $7/4^+$-power free ternary words with only one $7/4$-power is 102 letter long.

The two unavoidable $7/4$-powers occurring in the word are $(0121)^{7/4} = 0121012$ and $(2010)^{7/4} = 2010201$.
Up to a permutation of letters.

Morphism

g from $\{\mathtt{a},\mathtt{b},\mathtt{c},\mathtt{d}\}^*$ to $\{0,1,2\}^*$ defined by:

$$
\begin{cases}
g(\mathtt{a}) = & 0102101202102010210121020120210120102120121020120210121 \\
& 0212010210121020102120121020120210121020102101202102012 \\
& 10212010210121020120210120102120121020102101210212, \\
g(\mathtt{b}) = & 0102101202102010210121020120210120102120121020120210121 \\
& 0201021012021020121021201021012102010212012102012021012 \\
& 10212010210121020120210120102120121020102101210212, \\
g(\mathtt{c}) = & 0102101202102010210121020120210120102120121020102101202 \\
& 1020121021201021012102010212012102012021012102010210120 \\
& 21020102120121020120210120102120121020102101210212, \\
g(\mathtt{d}) = & 0102101202102010210121020120210120102120121020102101202 \\
& 1020102120121020120210121020102101202102012102120102101 \\
& 21020102120121020120210120102120121020102101210212.
\end{cases}
$$

The morphism is uniform with codeword length 160. The morphism g translates any infinite $7/5^+$-free word on the alphabet a, b, c, d into a $7/4^+$-free ternary word containing only two $7/4$-powers, the fewest possible.

Proof based on Ochem's result

Ochem's Lemma

Let $\alpha, \beta \in \mathbb{Q}$, $1 < \alpha < \beta < 2$, and $p \in \mathbb{N}^*$. Let $h : \Sigma_s^* \to \Sigma_e^*$ be a synchronizing q-uniform morphism (with $q \geq 1$). If $h(w)$ is (β^+, p)-free for every α^+-free word w such that $|w| < \max\{\frac{2\beta}{\beta-\alpha}, \frac{2(q-1)(2\beta-1)}{q(\beta-1)}\}$, then $h(t)$ is (β^+, p)-free for every (finite or infinite) α^+-free word t.

Application to morphism g

$\alpha = 7/5$ and $q = 160$.
Choose $\beta = 17/10$ and $n = 5$.
$g(w)$ is $(17/10^+, 5)$-power free for all words w for which

$$|w| < \max\{\frac{2\beta}{\beta - \alpha}, \frac{2(q-1)(2\beta-1)}{q(\beta-1)}\},$$

which implies $|w| < 12$.

An infinite ternary word with only one square

The morphism h from $\{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}^*$ to $\{0, 1, 2\}^*$ defined by:

$$\begin{cases} h(\mathtt{a}) = & 010020102120 \\ h(\mathtt{b}) = & 010021020120 \\ h(\mathtt{c}) = & 021001201020 \\ h(\mathtt{d}) = & 021012021020. \end{cases}$$

The morphism is uniform with codeword length 12.

The morphism h translates any infinite $7/5^+$-free word on the alphabet $\{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$ into a overlap-free ternary word containing only one square, the fewest possible.

Conjecture and Questions

Conjecture

The *finite-repetitions threshold* of 4-letter alphabet is $\frac{7}{5}$ and the associated number of $\frac{7}{5}$-powers is 2.

List of Repetitions: $\{(acdbcadcbd)^{\frac{7}{5}}, (bcdacbdcad)^{\frac{7}{5}}\}$,

Computation shows that the maximal length of $(7/5)^+$ free 4-ary word with only one 7/5-repetition is 230.

Conjecture tested up to length 100000.

Questions

- Finite Repetition Threshold for larger alphabet sizes?