Ultimately periodic and n-divided words

Giuseppe Pirillo
CNR- IAMI Sez. Firenze
viale Morgagni 67/A
I-50134 FIRENZE
ITALIA

and

Stefano Varricchio
Dipartimento di Matematica - Università
via Vetoio (loc. Coppito)
I-67100 L'AQUILA
ITALIA

Abstract. *We prove that if an infinite word does not contains n-divided factors then it is ultimately periodic.*

Our terminology is that usual in theoretical computer science [1]. In particular, the *free monoid* (resp. *free semigroup*) generated by the *alphabet* A is denoted by A^* (resp. A^+). We call the elements of A^* (finite) *words* and those of A *letters* and we denote by $|u|$ the length of a word u of A^*.

We denote by N the set of the non-negative integers and we extend the notion of a word to infinite words. A (right) *infinite word* on A is a map t from N into A. We write

$$t=t(0)t(1)\ldots t(i)\ldots$$

We say that a word u is a k-power if there exists a word v such that $u=v^k$.

Let t be a right-infinite word and p be a positive integer. We say that t is *ultimately periodic of period* p (in short *ultimately* p*-periodic*) if for some i_0 in N we have $t(i+p)=t(i)$ for $i\geq i_0$.

Suppose now that A is endowed whith a *total order* and consider on A^+ the *lexicographic order* induced by it. If $u,v\in A^+$ we write $u<v$ if u *strictly precedes* v in this order.

Let $x_1x_2\ldots x_n$ be a factorization of the word x and let σ be a non trivial element of the *symmetric group* Σ_n. We write x_σ for

$$x_{\sigma(1)}x_{\sigma(2)}\ldots x_{\sigma(n)}.$$

Definition 1. *A word* x *is* n*-divided if it admits an* n*-divided factorization* $x_1x_2\ldots x_n$, i.e. *a factorization such that for each* $\sigma\in\Sigma_n$*-{id} one has*

$$x>x_\sigma.$$

Definition 2. *An infinite word* t *is ultimately* ω*-divided if it admits a factorization*

$$t=t_0t_1t_2\ldots t_i\ldots$$

such that for each $i\in N$*-{0} and for each* $n\geq 2$

$$t_i\ldots t_{n+i-1}$$

is an n*-divided factorization.*

The following is a version of a famous theorem of Shirshov.

Theorem 1. (Shirshov, [1]). *Let $k,r,n\geq 1$ be integer such that $r\geq 2n$. There exist an integer $N(k,r,n)$ such that for any totally ordered alphabet A with k letters, any word in A^+ of length $N(k,r,n)$ contains as a factor either an n-divided word or a r-power, say u^r, with $0<|u|<n$.*

Recently several papers has appeared on this subject, for example [2-7]. In particular in [5] is proven the following result: *given any finite or infinite alphabet A and a total order on it then each infinite word on A is either ultimately periodic either ultimately ω-divided for the given total order either ultimately ω-divided for the inverse of the given total order.*

The aim of this paper is to prove the following theorem:

Theorem 2. *Let s be an infinite word on a finite alphabet A and n be an integer greater than 1. If there exist a total order on A such that s does not contain an n-divided factor then there exist a positive integer $p\leq n$ such that s is ultimately p-periodic.*

Proof. By way of contradiction, suppose that s is not ultimately p-periodic.

By Theorem 1, for each $\rho\geq 2n$, there exists a factor w_ρ of s such that

$$|w_\rho|\leq n$$

and that

$$(w_\rho)^\rho$$

is a factor of s.

As A is finite, there exists an infinite subset R of N and a word u such that if ρ is in R then the $w_\rho=u$.

So u^k is a factor of s for each $k\geq 1$.

We claim that we can factorise s in the following way:

$$s=r_1u^{i(1)}t_1r_2u^{i(2)}t_2\ldots r_{h-1}u^{i(h-1)}t_{h-1}r_hu^{(h)}t_h\ldots$$

with

$$|u|=|t_1|=|t_2|=\ldots=|t_{h-1}|=|t_h|=\ldots,$$
$$u\neq t_m\ ,$$

for each $m\geq 1$, and

$$1\leq i(1)<i(2)<\ldots<i(h-1)<i(h)<\ldots.$$

Let r_1 the shortest left factor of s such that for a suitable infinite word $s^{(1)}$, one has $s=r_1s^{(1)}$ and u is a left factor of $s^{(1)}$. Clearly r_1 exist. Let $i(1)$ the greatest integer such that $u^{i(1)}$ is a left factor of $s^{(1)}$. The integer $i(1)$ exist otherwise s must be ultimately p-periodic. Clearly $1\leq i(1)$ and $r_1u^{i(1)}$ is a left factor of s. Let t_1 the word of length $|u|$ such that $r_1u^{i(1)}t_1$ is a left factor of s. By maximality of $i(1)$ we have that u is different from t_1.

Now suppose that, for $h\geq 2$,

$$r_1u^{i(1)}t_1r_2u^{i(2)}t_2\ldots r_{h-1}u^{i(h-1)}t_{h-1}$$

is a left factor of s such that

$$|u|=|t_1|=|t_2|=\ldots=|t_{h-1}|,$$
$$1\leq i(1)<i(2)<\ldots<i(h-1)$$

and for each i, $1\leq i\leq h-1$, t_i is different from u.

Pose

$$s=r_1u^{i(1)}t_1r_2u^{i(2)}t_2\ldots r_{h-1}u^{i(h-1)}t_{h-1}s^{(h)}$$

where $s^{(h)}$ is a suitable infinite word.

Let r_h be the shortest left factor of $s^{(h)}$ such that for a suitable infinite word $s^{(h+1)}$

$$s^{(h)}=r_hs^{(h+1)}$$

and

$$u^{i(h-1)+1}$$

is a left factor of $s^{(h+1)}$.

The word r_h exist becouse u^k is a factor of s for each $k\geq 1$.

Let $i(h)$ be the greatest integer such that $u^{i(h)}$ is a left factor of $s^{(h+1)}$. The integer $i(h)$ exist otherwise s must be ultimately p-periodic. Let t_h be the word of length $|u|$ such that

$$r_1u^{i(1)}t_1r_2u^{i(2)}t_2\ldots r_{h-1}u^{i(h-1)}t_{h-1}r_hu^{(h)}t_h$$

is again a left factor of s.

Clearly

$$|u|=|t_1|=|t_2|=\ldots=|t_{h-1}|=|t_h|$$

One has again

$$1\leq i(1)<i(2)<\ldots<i(h-1)<i(h)$$

becouse $i(h)\geq i(h-1)+1>i(h-1)$.

Finally also t_h is different from u by maximality of $i(h)$.

This complete the proof of the claim.

Now, as A is finite there exist an infinite subset J of N and a word v such that, for each j in J, $t_j=v$ (and hence $|u|=|v|$). Let $j_1,j_2,\ldots,j_n,j_{n+1},j_{n+2},\ldots,j_{2n-1}$ elements of J such that

$$j_1<j_2<\ldots<j_n<j_{n+1}<j_{n+2}<\ldots<j_{2n-1}$$

We have two cases to consider.

Case $v>u$. The word

$$(u^{i(j_1)}v\ldots r_{j_2})(u^{i(j_2)}v\ldots r_{j_3})\ldots(u^{i(j_n)}v)$$

is clearly n-divided. Contradiction.

Case $u>v$. Remark that $i(j_{n+m})\geq n-m$ for each m, $0\leq m\leq n-1$. So we can pose $u^{i(j_{n+m})}=z'_{j_{n+m}}z_{j_{n+m}}$, where $|z_{j_{n+m}}|=|u|^{n-m}$. The word

$$(z_{j_n}v\ldots z'_{j_{n+1}})(z_{j_{n+1}}v\ldots z'_{j_{n+2}})\ldots(z_{j_{2n-1}}v)$$

is clearly n-divided. Contradiction.■

REFERENCES

1. M. LOTHAIRE, *Combinatorics on words*, Encyclopedia of Math. and its Applications, n.17, Addison-Wesley, London (1983).
2. A. de LUCA, S. VARRICCHIO, *Combinatorial Properties of Uniformly Recurrent Words and an Application to Semigroups*, International Journal of Algebra and Computation, Vol. 1, N.2, 1991.
3. J. JUSTIN, G. PIRILLO, *Shirshov's theorem and ω-permutability of semigroups*, Advances in Math. 87,2 (1991) 151-159.
4. J. JUSTIN, G. PIRILLO, *Factorial languages and some combinatorial properties of semigroups*, LITP, 1991.
5. G. PIRILLO, *Sur un théorème combinatoire de Shirshov*, C. R. Acad. Sci. Paris 313,1 (1991) 631-634.
6. A. RESTIVO, C. REUTENAUER, *Some applications of a theorem of Shirshov to language theory*, Information and Control 57 (1983) 205-213.
7. S. VARRICCHIO, *Factorizations of Free Monoids and Unavoidable Regularities*, Theoret. Comput. Sci., 73, 81-89 (1990).

Une méthode pour obtenir la fonction génératrice algébrique d'une série.

Simon Plouffe
LaBRI, Université Bordeaux I
351, cours de la Libération
Talence 33405
France
Mars 1993.

Résumé

Nous décrivons ici une méthode expérimentale permettant de calculer de bons candidats pour une forme close de fonctions génératrices à partir des premiers termes d'une suite de nombres rationnels. La méthode est basée sur l'algorithme LLL[1] et utilise deux programmes de calcul symbolique, soit MapleV et Pari-GP. Quelques résultats sont présentés en appendice. Cette méthode a été testée sur toute la table de suites du livre , *The New book of Integer Sequences*, de N.J.A Sloane et S. Plouffe (en préparation). Ainsi, nous avons obtenu de cette façon la fonction génératrice,

$$\frac{z + (z+1)^{1/2}\,(1-3z)^{1/2} - 1}{2\left(z^2\,(z+1)^{1/2}\,(1-3z)^{1/2}\right)}$$

pour la suite: 1, 2, 6, 16, 45, 126, 357, 1016, 2907, 8350, 24068, 69576, 201643, 585690,... qui apparaît en page 78 du livre de Louis Comtet, *Adanced Combinatorics*.

[1] Nommé ainsi à cause des travaux de Lenstra, Lenstra et Lovasz.