ADVANCES IN MATHEMATICS 87, 151-159 (1991)

Shirshov's Theorem and ω-Permutability of Semigroups

J. JUSTIN

LITP-CNRS-Université Paris VII, 19, rue de Bagneux, 92330 Sceaux, France

AND

G. PIRILLO

IAGA-IAMI CNR, Viale Morgagni, 67/A, 50134 Florence, Italy

DEDICATED TO MARCEL-PAUL SCHÜTZENBERGER

1. INTRODUCTION

A theorem of Shirshov [11] states that each long enough word over a finite alphabet contains either a factor which is "n-divided" or a factor which is a pth power.

Reutenauer [12] has given an elegant proof of this result using properties of Lyndon's words. It is remarkable that Shirshov's theorem can be generalized and extended to right infinite words; this was shown by Varricchio [14]. Inspired by the latter work and making use of the properties of uniformly recurrent infinite words [3], we are able to provide a very simple and self-contained proof, which is also valid for two-sided infinite words (a case which does not seem to be easily tractable using Varricchio's method).

Blyth and Rhemtulla [1] and de Luca and Varricchio [2] have introduced the notion of right ω-permutability of a semigroup. It has been shown in [2] that this property is weaker than permutability; a further complication of the method allows us to prove that the right, left, and two-sided ω-permutabilities are distinct. On the other hand, the three properties of weak ω-permutability are equivalent.

The link between permutability and Shirshov's theorem is Restivo and Reutenauer's theorem [10], stating that a finitely generated periodic and permutable semigroup is finite. This theorem has been extended in [2] to

0001-8708/91 $7.50

the right ω-permutability (or left, according to duality). We similarly extend it to the two-sided ω-permutability.

Finally, we consider how much certain preservation properties pertaining to permutability [5, 6] can be extended to the ω-permutabilities, using an interesting property of finite partitions of the free semigroup, the ω-repetitivity, which is a corollary of Ramsey's theorem in infinite version and was first noticed (and directly proved) by Schützenberger [4, 13].

2. Shirshov's Theorem

We refer to [7] for the terminology concerning the free monoid A^* and the free semigroup $A^+ = A^* - \{1\}$ generated by the alphabet A. We extend the notion of a word to infinite words: a right (resp. left, resp. two-sided) infinite word over A is a map t of $\mathbb{N}$ (resp. $-\mathbb{N}$, resp. $\mathbb{Z}$) into A.

If t is a word (either finite or infinite), let $t(i)$ be the letter of A occurring at "rank i" in t and let $t(i, j)$, $i \leqslant j$, be the factor $t(i) \cdots t(j)$ of t. By word and factor we will always mean a finite nonempty word, excepted where otherwise stated. If E is a set of words (either finite or infinite) we will denote by $F(E)$ the set of the factors of the elements of E. If $E = \{m\}$, where m is a word (either finite or infinite), we will simply write $F(m)$.

The proof of the following lemma makes use of a construction whose well-known principle goes back to some lemmas of König and Rado.

Lemma 1. *Let A be a finite alphabet. If $E \subset A^+$ is infinite, there exists a two-sided infinite word b such that each factor of b is a factor of infinitely many elements of E. In particular, if s is an infinite word over A, there exists a two-sided (and, a fortiori, right or left) infinite word b such that $F(b) \subset F(s)$.*

Proof. For each $u \in E$ such that $|u| \geqslant 2$ let $r_u = [|u|/2]$.

There are infinitely many u's in E such that the same letter a_0, say, occurs at rank r_u in u. Infinitely many of them satisfy

$$u(r_u, r_u + 1) = a_0 a_1$$

for some letter a_1. In the same way there exist a letter a_2 and infinitely many words u such that

$$u(r_u - 1, r_u + 1) = a_2 a_0 a_1.$$

Going on like this, we construct the two-sided infinite word

$$\cdots a_4 a_2 a_0 a_1 a_3 \cdots$$

in which each factor is a factor of infinitely many words of E. ∎

Let s be an infinite word and $u \in F(s)$. We will write

$$g(s, u) = \text{Sup} \{ |w| ; w \in F(s), u \notin F(w) \}.$$

In particular $g(s, u) = \infty$ if s has arbitrarily long factors where u does not occur.

DEFINITION. An infinite word t is uniformly recurrent if for each factor u of t one has $g(t, u) < \infty$.

LEMMA 2. *If s is an infinite word over a finite alphabet there exists a uniformly recurrent two-sided infinite word b such that $F(b) \subset F(s)$.*

Proof. Let $u_1, u_2, \ldots,$ be an arbitrary enumeration of the factors of s. We construct an infinite chain of infinite words in the following way.

Let $t_0 = s$.

For each $i > 1$, if $u_i \in F(t_{i-1})$ and $g(t_{i-1}, u_i) < \infty$ we let $t_i = t_{i-1}$. In the opposite case, let E be the infinite set of the factors v of t_{i-1} such that $u_i \notin F(v)$. We construct t_i from the set E using Lemma 1.

From this construction it turns out that if a word w is factor of a certain t_j and if $w = u_k$ with $j \geqslant k$, then $g(t_j, w) \leqslant g(t_{k-1}, w) < \infty$.

Let us now choose, in each t_i, $i > 0$, a factor v_i having length i and let E be the infinite set of these factors. According to Lemma 1 there exists an infinite two-sided word b such that each factor of b is a factor of infinitely many words in E.

Therefore, if $w = u_k \in F(b)$, we have $w \in F(t_j)$ for some j greater than k. Hence, by above, $g(t_j, w) < \infty$. Finally, $g(b, w) < \infty$ since $F(b) \subset F(t_j)$. Thus, b is uniformly recurrent. ∎

DEFINITIONS. A right (resp. left) or two-sided infinite word t is ultimately periodic on the right (resp. on the left) if there exist $k > 0$ and $i_0 \in \mathbb{Z}$ such that for each $i \geqslant i_0$ (resp. for each $i \leqslant i_0$) one has $t(i+k) = t(i)$.

An infinite word t is periodic if there exists $k > 0$ such that $t(i+k) = t(i)$ for each i and $i+k$ in the domain of t.

It is obvious that if a uniformly recurrent word is ultimately periodic on one side it is periodic.

LEMMA 3. *Let t be a uniformly recurrent two-sided infinite word and let $w \in F(t)$. If t is not periodic there exist two different words, u and v, having the same length, such that $wu, wv \in F(t)$.*

Proof. Let $t(m, n)$ and $t(m+k, n+k)$ be two different occurrences of w

in t. If the conclusion of the lemma were not satisfied we would have for each $i \geqslant m$

$$t(m, i) = t(m+k, i+k),$$

hence, in particular

$$t(i) = t(i+k).$$

Therefore t would be periodic, which is a contradiction. ∎

From now on we will suppose that A^+ is endowed with a lexicographical order and we will write $u < v$ if u strictly precedes v according to this order.

DEFINITION. The strong lexicographical relation, denoted $\ll$, is defined by $u \ll v$ if and only if $u = fx$, $v = gy$ with $f, g \in A^+$, $x, y \in A^*$, $|f| = |g|$, and $f < g$.

It is clear that $\ll$ is transitive and that for each $x, y \in A^*$

$$u \ll v \text{ implies } xu \ll xv \text{ and } ux \ll vy.$$

LEMMA 4. *Let t be a uniformly recurrent and nonperiodic two-sided infinite word. There exist infinitely many factors of t, say w_i ($i \in \mathbb{Z}$), such that*

$$w_i \gg w_j$$

if $i < j$.

Proof. We inductively define the w_j's by means of the recurrence hypothesis:

H(i). The w_j's have been defined for $-i \leqslant j \leqslant i$ and there exist $x, y \in F(t)$ such that $x \gg w_{-i}$ and $w_i \gg y$.

As t is not periodic, there exist, by Lemma 3, two words u and v such that $u \gg v$ and $yu, yv \in F(t)$. Let $w_{i+1} = yu$ and $y' = yv$. We have $w_i \gg w_{i+1} \gg y'$.

In the same way there exist two words e and f such that $e \gg f$ and $xe, xf \in F(t)$. Let $w_{-i-1} = xf$ and $x' = xe$. We have $x' \gg w_{-i-1} \gg w_{-i}$.

Hypothesis H($i+1$) is therefore verified. We verify Hypothesis H(0) in a similar way. ∎

LEMMA 5. *Let t be a uniformly recurrent and nonperiodic two-sided infinite word. Then there exists a factorization*

$$t = \cdots u_{-n} \cdots u_0 \cdots u_n \cdots$$

such that

$$u_i \gg u_j$$

if $i<j$.

Proof. Let the w_i's be as in Lemma 4. There exist (possibly empty) words, f_i, $i \in \mathbb{Z}$, such that t can be factorized as

$$\cdots w_{-n} f_{-n} \cdots w_0 f_0 \cdots w_n f_n \cdots.$$

Letting $u_i = w_i f_i$ for $i \in \mathbb{Z}$, we obtain the wanted factorization. ∎

Now let $x_1 x_2 \cdots x_n$ be a factorization of a word x and σ be an element of the symmetric group $\sum_n$. We write x_σ for $x_{\sigma(1)} \cdots x_{\sigma(n)}$. Let us recall the following definition.

DEFINITION. A word x is n-divided if it admits an n-divided $x_1 \cdots x_n$ factorization, i.e., a factorization such that for each $\sigma \in \Sigma_n$-id we have $x > x_\sigma$.

LEMMA 6. *If* $x_1 \gg x_2 \gg \cdots \gg x_n$ *then* $x = x_1 x_2 \cdots x_n$ *is an n-divided factorization.*

Proof. Let i be minimal such that $\sigma(i) \neq i$. As $\sigma(i) > i$ we have $x_{\sigma(i)} \ll x_i$, hence $x_\sigma < x$. ∎

DEFINITION. A two-sided infinite word is ω-divided if it admits a factorization $\cdots w_i w_{i+1} \cdots$ such that, for each $i \in \mathbb{Z}$ and for each $n > 0$, $w_i w_{i+1} \cdots w_{n+i-1}$ is an n-divided factorization.

The following generalization of Shirshov's theorem obviously follows from Lemmas 2, 5, and 6.

THEOREM 1. *If s is an infinite word over a finite alphabet A and if A^+ is endowed with a lexicographical order, there exists a two-sided infinite word t such that $F(t) \subset F(s)$ and t is periodic or ω-divided.*

3. THE ω-PERMUTABILITY

Let $x_1, x_2, \ldots, x_n$ ($n \geq 2$) be elements of a semigroup S. We say that the product $x = x_1 x_2 \cdots x_n$ is permutable (resp. weakly permutable) if there exists $\sigma \in \Sigma_n$-id such that $x_\sigma = x$ (resp. if there exist $\sigma, \tau \in \Sigma_n$, $\sigma \neq \tau$, such that $x_\sigma = x_\tau$).

The semigroup S is said to be permutable (resp. weakly permutable) if there exists $n > 1$ such that each product of n elements of S is permutable (weakly permutable). A weakly permutable semigroup (even if finitely generated) is not necessarily permutable [5, 8, 9].

DEFINITION. The semigroup S is right (resp. left, resp. two-sided) ω-permutable if in each word s over the alphabet S which is right (resp. left, resp. two-sided) infinite there exists a factor, say $s(m) \cdots s(n)$, which is permutable as a product in S.

We define the weak properties in a similar way.

Let P, $P\omega R$, $P\omega L$, and $P\omega B$ denote the permutability and the three ω-permutabilities and let P^*, $P^*\omega R$, $P^*\omega L$ and $P^*\omega B$ denote the corresponding weak properties.

Theorem 1 allows one to establish (just with the same methods as those in [2, 10]) the following theorem.

THEOREM 2. *A finitely generated periodic semigroup is two-sided ω-permutable if and only if it is finite.*

This theorem is of some interest only if $P\omega B$ is different from $P\omega R$.

This is what we shall prove.

Let A be the alphabet $\{a_i;\ i \in \mathbb{N}\}$ and, for $i \in \mathbb{N}$, let A_i be the set of the words $a_{i_1} a_{i_2} \cdots a_{i_r}$ such that $i_1 = \mathrm{Inf}\{i_1, i_2, ..., i_r\}$ and $i_1 = i$.

Each word w can be factorized in a unique way in the form

$$w = u_1 u_2 \cdots u_n$$

with

$$u_1 \in A_{j_1}, ..., u_n \in A_{j_n}$$

and

$$j_1 > j_2 > \cdots > j_n.$$

Let E be the set of the words w, so factorized, such that

$$j_1 > |u_1|, j_2 > |u_2|, ..., j_n > |u_n|.$$

We have $F(E) = E$. It is easy to see that the Rees quotient of A^+ by the ideal $A^+ - E$ has $P\omega R$ but does not have $P\omega L$. In fact each right infinite word over A has some left factors belonging to the ideal. On the other hand, none of the factors of the left infinite word

$$\cdots a_4 a_3 a_2$$

belongs to the ideal.

Therefore $P\omega R$ does not imply $P\omega L$ (and vice versa). As each of them implies $P\omega B$, the three properties are different.

Remark. It is even possible to construct semigroups which have neither $P\omega R$ nor $P\omega L$ but have $P\omega B$.

On the other hand $P^*\omega B$, $P^*\omega R$, and $P^*\omega L$ are equivalent. To see this, assume that S has $P^*\omega B$ and let $d = s_0 s_1 s_2 \ldots$, $s_i \in S$, be a right infinite word over the alphabet S.

The two-sided infinite word $\cdots s_2 s_0 s_1 s_3 \cdots$ contains a weakly permutable factor $s_i \cdots s_j$. Then d contains a weakly permutable factor $s_0 \cdots s_n$ with $n = \operatorname{Sup}\{i, j\}$. On the other hand P^* and $P^*\omega$ (a common notation for $P^*\omega B$, $P^*\omega L$, and $P^*\omega R$) are different, as appears from the example given in [2] for P and $P\omega R$.

The nonequivalence of P with $P\omega R$ and the nonequivalence between $P\omega R$, $P\omega L$, and $P\omega B$, proved for non-finitely-generated semigroups, remain true in the case of finitely generated semigroups. It suffices for seeing that to code the countable alphabet $\{a_i;\ i \in \mathbb{N}\}$ by means of an alphabet $X = \{x, y\}$, letting xy^i correspond to a_i, and to take care in the definition of the ideal.

More precisely, let X_α be the set of the words having the form

$$xy^\alpha xy^\beta \cdots xy^\lambda$$

with $\alpha, \beta, \ldots, \lambda \geqslant 0$ and $\alpha = \operatorname{Inf}\{\alpha, \beta, \ldots, \lambda\}$.

Each word of X^+ can be uniquely factorized in the form

$$y^n v_{i_1} \cdots v_{i_t}$$

with $n \geqslant 0$, $t \geqslant 0$, $v_{i_1} \in X_{i_1}, \ldots, v_{i_t} \in X_{i_t}$, and

$$i_1 > i_2 > \cdots > i_t.$$

Let K be the set of the words, so factorized, such that

$$i_1 + 2 > |v_{i_1}|_x, \ldots, i_t + 2 > |v_{i_t}|_x$$

(where $|w|_x$ represents the number of occurrences of x in w).

We easily verify that $F(K) = K$ and that the Rees quotient of X^+ by the ideal $X^+ - K$ has $P\omega R$ and does not have $P\omega L$.

Remark. In the case of groups $P\omega R$, $P\omega L$ and $P\omega B$ are equivalent. Indeed, $P\omega R$ is shown in [1] to be equivalent to P. By very slight alterations in the proof it appears that $P\omega B$ is also equivalent to P.

For the end we now state four preservation theorems for $P\omega R$ (and $P\omega L$, by duality) and $P^*\omega$.

The proofs are similar to those of the corresponding theorems for P and P^*. In particular, instead of using the repetitivity of the finite partitions of the free semigroups (corollary of Ramsey's theorem) [4], we use, in order to prove Theorems 3 and 4, the ω-repetitivity of these same partitions [4, 13], which is a corollary of Ramsey's theorem in infinite version and which can be stated as follows: let α be a map from A^+ into a finite set; then each right infinite word over A can be factorized in the form $w_0 w_1 w_2 \ldots$, where all the w_i's, $i \geqslant 1$, have the same image under α.

THEOREM 3. *Let $\alpha: S \to F$ be a surjective morphism from a semigroup S onto a finite semigroup F. If for each idempotent e of F, $\alpha^{-1}(e)$ has $P\omega R$ (resp. $P^*\omega$), then the same holds for S.*

THEOREM 4. *If the semigroup S is the union of a finite number of subsemigroups and if each of them has $P\omega R$ (resp. $P^*\omega$), then the same holds for S.*

THEOREM 5. *Let I be a two-sided ideal of the semigroup S. If I and the Rees quotient S/I have $P\omega R$ (resp. $P^*\omega$), then the same holds for S.*

THEOREM 6. *Let $\alpha: S \to D$ be a morphism of semigroups such that for some positive integer m one has for each $d \in D$, $\mathrm{Card}(\alpha^{-1}(d)) < m$. If D has $P^*\omega$, then the same holds for S.*

Remark. We do not know whether Theorems 3, 4, and 5 remain true when one replaces $P\omega R$ by $P\omega B$.

In particular, as far as Theorems 3 and 4 are concerned, if they remain true, it is not possible to prove them by using an extension to the two-sided case of the ω-repetitivity of the finite partitions, because such an extension is not possible, as the following example shows.

Let $A = \{a, b\}$ and $G = \{0, 1\}$ be the additive group of the integers modulo 2 and $\alpha: A^+ \to G$ be the morphism given by $\alpha(a) = 0$ and $\alpha(b) = 1$. Then the two-sided infinite word

$$\cdots a \cdots aba \cdots a \cdots$$

cannot be factorized in the form

$$\cdots w_{i-1} w_i w_{i+1} \cdots,$$

where all the w_i's have the same image under α.

Remark. A French version of this paper has appeared as the Technical Report LITP-Univ. Paris VII No. 22/89 (March 1989) under the title "Théorème de Shirshov et ω-permutabilité des semi-groupes."

REFERENCES

1. R. D. BLYTH AND A. H. RHEMTULLA, Rewritable products in FC-by-finite groups, *Canad. J. Math.* **26** (1989), 369–384.
2. A. DE LUCA AND S. VARRICCHIO, A note on ω-permutable semigroups, *Semigroup Forum* **40** (1990), 153–157.
3. N. FÜRSTENBERG, "Recurrence in Ergodic Theory and Combinatorial Number Theory," Princeton Univ. Press, Princeton, NJ, 1981.
4. J. JUSTIN AND G. PIRILLO, On a natural extension of Jacob's ranks, *J. Combin. Theory Ser. A* **43**, No. 2 (1986), 205–218.
5. J. JUSTIN AND G. PIRILLO, Comments on the permutation property for semigroups, *Semigroup Forum* **39**, No. 1 (1989), 109–112.
6. J. JUSTIN AND G. PIRILLO, Some remarks on the permutation property for semigroups, *European J. Combin.* **11** (1990), 151–154.
7. D. PERRIN, Chapter 1 *in* M. Lothaire, "Combinatorics on Words," Addison–Wesley, Reading, MA, 1984.
8. G. PIRILLO, On a permutation property for semigroups, *in* "Group Theory Conference, Bressanone/Brixen, 1986," Lecture Notes in Mathematics, Vol. 1281, Springer-Verlag, Berlin, 1987.
9. G. PIRILLO, On permutation properties for finitely generated semigroups, *in* "Proceedings, Combinatorics '86, Passo della Mendola, 1986"; *Ann. Discrete Math.* **37** (1988).
10. A. RESTIVO AND C. REUTENAUER, On the Burnside problem for semigroups, *J. Algebra* **89** (1984) 102–104.
11. C. REUTENAUER, Chapter 7 *in* M. Lothaire, "Combinatorics on Words," Addison–Wesley, Reading, MA, 1984.
12. C. REUTENAUER, Mots de Lyndon et théorème de Shirshov, *Ann. Sci. Math.* **10**, No. 2 (1986), 237–245.
13. M. P. SCHÜTZENBERGER, "Quelques problèmes combinatoires de la théorie des automates" (J. F. Perrot, Ed.), cours, Institut de Programmation, Univ. Paris (1966/1967).
14. S. VARRICCHIO, Factorizations of free monoids and unavoidable regularities, *Theoret. Comput. Sci.* **73** (1990), 81–89.