Available online at www.sciencedirect.com
ScienceDirect
Theoretical Computer Science 366 (2006) 194–198
Theoretical Computer Science
www.elsevier.com/locate/tcs

Alternative proof of the Lyndon–Schützenberger Theorem

Pál Dömösi*, Géza Horváth

Faculty of Informatics, Debrecen University, Egyetem tér 1., Debrecen H-4032, Hungary

To the memory of Professor Alexandru Mateescu

Abstract

Some observations on products of primitive words are discussed. By these results, alternative proof is given for the Lyndon–Schützenberger Theorem, which says that every solution of the equation $a^m b^n = c^k$ over Σ^* is trivial.

Keywords: Combinatorics of words

1. Introduction

A word is primitive if it is not empty and not a power of another word. A well-known unsolved problem is in theoretical computer science whether the language of all primitive words over a nontrivial alphabet is context free [4,5]. Among others, this (in)famous problem motivates the study of combinatorial properties of primitive words. In addition, they have special importance in studying automatic sequences [1,9]. The Lyndon–Schützenberger Theorem is a well-known classical result in this direction. The aim of this paper is to give alternative proof of this celebrated theorem.

Some of the known proofs of this famous result is rather involved [2,10–12]. On the other hand, the Lyndon–Schützenberger Theorem also has simple proofs, see [7,9]. We give a proof which is different in the technical details.

We note that the original form of the Lyndon–Schützenberger Theorem was proved for free groups in [10].

In our alternative proof of the Lyndon–Schützenberger Theorem, we follow the structures of the proofs in [2,7,9,11,12]. For the sake of completeness, we also describe the proof of the "easy" case (*Case* 2), which is essentially the same as the corresponding proof in [2,7,9,11,12]. The alternative proof of the "difficult" case (*Case* 1) is on the basis of new observations which cannot be found in the cited works.

2. Preliminaries

By an *alphabet* we mean a finite, nonempty set Σ, the elements of which are called letters. Σ is called *trivial* if it is a singleton. Otherwise we also say that Σ is *nontrivial*. A *word* over Σ is a finite sequence of elements of Σ. If there is

* Corresponding author.
E-mail addresses: domosi@inf.unideb.hu (P. Dömösi), geza@inf.unideb.hu (G. Horváth).

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2006.08.023

no danger of confusion, sometimes we omit the expression "over Σ". We also define the *empty word* λ consisting of zero letters. Given a word $w = x_1, \ldots, x_n$ with $x_1, \ldots, x_n \in \Sigma$, we put $w^R = x_n, \ldots, x_1$, i.e. w^R denotes the *mirror image* of w. In addition, we put $\lambda^R = \lambda$. The set of all words over Σ is denoted by Σ^* as usual. Moreover, we put $\Sigma^+ = \Sigma^* \setminus \{\lambda\}$. Note that Σ^*, equipped with the operation *catenation*, is the *free monoid* generated by Σ, while Σ^+, equipped with the same operation, is the *free semigroup* over Σ. The catenation is also called *product*. The *length* $|w|$ of a word w is the number of letters in w, where each letter is counted as many times as it occurs. Thus, $|\lambda| = 0$. Two words $u, v \in \Sigma^*$ are said to be *conjugates* if there exists a word $w \in \Sigma^*$ with $uw = wv$. In particular, a word z is called *overlapping* or *bordered* if there are $u, v, w \in \Sigma^+$ with $z = uw = wv$. Otherwise we say that z is *nonoverlapping* or *unbordered*.

The following statement obviously holds.

Proposition 1. *Every bordered word can be written in the form uvu for some $u \in \Sigma^+$, $v \in \Sigma^*$.*

Next we recall some results which we will use.

Lemma 2 (*Lyndon and Schützenberger [10]*). *The words $u, v \in \Sigma^*$ are conjugates if and only if there are words $p, q \in \Sigma^*$ with $u = pq$ and $v = qp$.*

By the above result, for all words $p, q \in \Sigma^*$, it is also said that pq and qp are conjugates. Given a word u, we define $u^0 = \lambda$, and for $n > 0$, $u^n = u^{n-1}u$. Moreover, we put $u^* = \{u^n : n \geqslant 0\}$ and $u^+ = \{u^n : n \geqslant 1\}$. Thus, u^n with $n \geqslant 0$ are the nth *powers* of u. The next result concerns words which are conjugates.

Lemma 3 (*Lyndon and Schützenberger [10]*). *Let $u, v \in \Sigma^+$ with $uv = vu$. There exists $w \in \Sigma^+$ with $u, v \in w^+$.*

Lemma 4 (*Lyndon and Schützenberger [10]*). *If $uv = vq, q, u \in \Sigma^+, v \in \Sigma^*$, then $u = wz$, $v = (wz)^k w$, $q = zw$ for some $w \in \Sigma^*$, $z \in \Sigma^+$ and $k \geqslant 0$.*

Given a list $c_1, \ldots, c_n$ of integers, let $gcd(c_1, \ldots, c_n)$ denote the greatest common divisor of $c_1, \ldots, c_n$.

Theorem 5 (*Fine-Wilf Theorem [6]*). *Let $u, v \in \Sigma^*$. There exists a $w \in \Sigma^+$ such that $u, v \in w^+$ if and only if there are $i, j \geqslant 0$ so that u^i and v^j have a common prefix (suffix) of length $|u| + |v| - gcd(|u|, |v|)$.*

A word $v \in \Sigma^*$ is *primitive* if $v \neq \lambda$ and there are no $w \in \Sigma^+$ and $n \geqslant 2$ such that $v = w^n$. The set of all primitive words over Σ will be denoted by $Q(\Sigma)$, or simply by Q if Σ is understood. A *lexicographic ordering* ϱ on Σ^* is an extension of a strict linear ordering τ on the alphabet Σ in the following way: for every $u, v \in \Sigma^*$, $u \varrho v$ if and only if either $v \in \{u\}\Sigma^+$ or $u = raw$, $v = rbz$ with $a \tau b$, $a, b \in \Sigma$, $r, w, z \in \Sigma^*$. Given a lexicographic ordering ϱ on Σ^*, let w be a primitive word which is minimal among its conjugates with respect to ϱ. Then w is called a *Lyndon word with respect to* ϱ, or in short, a *Lyndon word* if ϱ is understood.

The following statement is obvious.

Lemma 6. *Let ϱ be a lexicographic ordering on Σ^*. For every $u, v, w, z \in \Sigma^*$ we have the following properties.*
(i) *$u \varrho v$ if and only if $wu \varrho wv$;*
(ii) *if u is not a prefix of v, then $u \varrho v$ implies $uw \varrho vz$.*

Lemma 7 (*Shyr and Thierrin [13]*). *Let $u, v, w \in \Sigma^*$, $i \geqslant 1$. If $w^i = uv$, then there are $p, q \in \Sigma^*$ with $w = pq$ and $(qp)^i = vu$. Furthermore, $uv \in Q$ for some $u, v \in \Sigma^*$ if and only if $vu \in Q$.*

Lemma 8 (*Lyndon and Schützenberger [10]*). *If $u \neq \lambda$, then there exists a unique primitive word f and a unique integer $k \geqslant 1$ such that $u = f^k$.*

Let $u \neq \lambda$ and let f be a primitive word with an integer $k \geqslant 1$ having $u = f^k$. We let $\sqrt{u} = f$ and call f the *primitive root* of the word u. Let $a^m b^n = c^k$ be an equation over Σ^* such that $m, n, k \geqslant 2$. A solution $a, b, c \in \Sigma^*$ of the above equation is called *trivial* if there is a $w \in \Sigma^*$ such that $a, b, c \in w^*$.

The next result was shown for free groups in [10]. Since every free monoid can be embedded in a free group, the result is true on a free monoid too.

Theorem 9 (*Lyndon–Schützenberger Theorem [10]*). *Every solution of the equation* $a^m b^n = c^k$ *over* Σ^* *is trivial.*

3. Main results

To the completeness of the paper, we recall the proof of the next statement given in [8].

Proposition 10. *Lyndon words are unbordered.*

Proof. If there exists a bordered Lyndon word then, by Proposition 1, it can be written in the form uvu, $u, v \in \Sigma^+$, $u \neq v$. With respect to the lexicographic order "$\leqslant$" then $uvu \leqslant uuv$ and so $vu \leqslant uv$, when the common prefix is removed. This yields that $vuu \leqslant uvu$, a contradiction. □

Now we show a short proof of the next result stated in [7] without proof.

Proposition 11. *Let* $v \in \Sigma^+$ *be an arbitrary word. There are* $u \in \Sigma^+$, $k \geqslant 2$ *with* $|u| < |v|$ *so that* v *is a subword of* u^k, *if and only if*, v *is bordered.*

Proof. (1) If v is bordered then $v = pqp$, and $u = pq$, $k = 2$ are appropriate.

(2) If v is a subword of u^k then $u = u_1u_2u_3$, $u_1, u_3 \in \Sigma^*$, $u_2 \in \Sigma^+$ and $v = u_2u_3u^*u_1u_2$ or $v = u_3u^+u_1 = u_3u_1(u_2u_3u_1)^*u_2u_3u_1$ or $v = u_2u_3u^+u_1 = u_2u_3(u_1u_2u_3)^+u_1$. □

Lemma 12. *Let* $u, v \in Q$, *such that* $u^m = v^k w$ *for some* $k, m \geqslant 2$, *and* $w \in \Sigma^*$ *with* $|w| \leqslant |v|$. *Then exactly one of the following conditions holds*:

(i) $u = v$ *and* $w \in \{u, \lambda\}$;

(ii) $m = k = 2$ *and there are* $p, q \in \Sigma^+$, $s \geqslant 1$ *with* $\sqrt{p} \neq \sqrt{q}$, $u = (pq)^{s+1}p^2q$, $v = (pq)^{s+1}p$, $w = qp^2q$.

Proof. The conditions $u = v$ and $u^m = v^k w$ imply $w = v^{m-k}$. Therefore, by $|w| \leqslant |v|$ and $u = v$, $w \in \{u, \lambda\}$. Thus, it remains to prove that exactly one of the conditions $u = v$ and (ii) holds.

Then v^k is a prefix of u^m and $v^k w$ with $m, k \geqslant 2$. Therefore, by Theorem 5, we have (i) whenever $|u| + |v| \leqslant |v^k|$. Thus, we may assume $|u| + |v| > |v^k|$. By $k \geqslant 2$, this implies $|v| < |u|$. Then, by $|w| \leqslant |v| < |u|$, $(m-1)|u| < |u^m| - |w| = |v^k| < |u| + |v| < 2|u|$. Therefore, $m < 3$ (with $m \geqslant 2$), i.e. $m = 2$. In this case, $2|u| = |v^k| + |w|$ which, using $|v^k| < |u| + |v|$, leads to $|u| - |w| < |v|$, or in another form, $|u| < |v| + |w|$.

Thus, we reached $u^2 = v^k w$ with $|w| \leqslant |v| < |u| < |v| + |w|$. Therefore, taking into consideration $|w| < |u|$, $u = v^\ell v_1 = v_2 v^{k-\ell-1} w$ for some $v_1, v_2 \in \Sigma^*$, $\ell \geqslant 0$ with $v = v_1v_2$ and $v_2 \neq \lambda$. Therefore, by $|u| < |v| + |w|$, we get $k - \ell - 1 = 0$. Hence, $u = v^{k-1}v_1 = v_2 w$.

Observe that $v_1 = \lambda$ implies $u = v^{k-1}$, which is impossible. Therefore, by $|v_2| + |w| \leqslant 2|v|$ and $k \geqslant 2$, we obtain $k = 2$. By $u = v^{k-1}v_1 = v_2 w$, this means $u = vv_1 = v_1v_2v_1 = v_2w_1v_1$ with $w = w_1v_1$. Hence, $v_1v_2 = v_2w_1$. Applying Lemma 4, there are $p, q \in \Sigma^*$, $s \geqslant 0$ having $v_1 = pq$, $v_2 = (pq)^s p$, $w_1 = qp$. Hence, $u = (pq)^{s+1}p^2q$, $v = (pq)^{s+1}p$, $w = qp^2q$. On the other hand, by $|w| \leqslant |v|$, $s \geqslant 1$. In addition, $u, v \in Q$ implies $\sqrt{p} \neq \sqrt{q}$ and also $\lambda \notin \{p, q\}$. Then $u \neq v$ and $w \notin \{u, \lambda\}$ are also obvious. Therefore, (i) does not hold whenever (ii) holds and vice versa. □

Theorem 13. *Let* $u, v \in Q$, *such that* $u^m = v^k w$ *for some prefix* w *of* v *and* $k, m \geqslant 2$. *Then* $u = v$ *and* $w \in \{u, \lambda\}$.

Proof. If $m = n = 2$ does not hold then this statement is a direct consequence of Lemma 12. Suppose $m = n = 2$ and $u = v$. Then $u^2 = v^2 w = u^2 w$ implies $w = \lambda$. Otherwise we should consider $m = n = 2$ with $u \neq v$. Then, by (ii) of Lemma 12, w is not a prefix of v. □

Theorem 14. *Let* $u, v \in Q$, *such that* $u^m = wv^k$ *for some suffix* w *of* v *and* $k, m \geqslant 2$. *Then* $u = v$ *and* $w \in \{u, \lambda\}$.

Proof. If $u, v \in Q$, such that $u^m = wv^k$ for some suffix w of v and $k, m \geqslant 2$, then $u^R, v^R \in Q$, such that $(u^R)^m = (v^R)^k w^R$ for some prefix w^R of v^R. Applying Theorem 13, we have $u^R = v^R$ and $w^R \in \{u^R, \lambda\}$. Therefore, $u = v$ and $w \in \{u, \lambda\}$. □

Now we are ready to give an alternative proof of Lyndon–Schützenberger Theorem:

Proof of Theorem 9. If $\lambda \in \{a, b, c\}$, then our statement is trivial. Thus we can assume $a, b, c \in \Sigma^+$. Clearly, then we may also assume $a, b, c \in Q$ without any restriction. In addition, it is clear that our statement holds either $b = \sqrt{b} = \sqrt{c} = c$ or $a = \sqrt{a} = \sqrt{c} = c$. Thus, let $b = \sqrt{b} \neq \sqrt{c} = c$ and $a = \sqrt{a} \neq \sqrt{c} = c$.

Let $a^m = c^s c_1, b^n = c_2 c^{k-s-1}$ with $c_1, c_2 \in \Sigma^+, c = c_1 c_2$. Suppose $s > 1$. Then, applying Theorem 13, $a = c$ which leads to $a = b = c$, a contradiction. Suppose $k - s - 1 > 1$. Hence, by Theorem 14, $b = c$ leading to $a = b = c$ again. It remains to study the case $0 \leqslant s \leqslant 1$, $0 \leqslant k - s - 1 \leqslant 1$. Using this assumption, by $k \geqslant 2$, we obtain $k \in \{2, 3\}$ so that $k = 3$ implies $s = 1$. We distinguish the following two cases.

Case 1: $k = 3$ with $s = 1$.

Then $a^m = c_1 c_2 c_1$ and $b^n = c_2 c_1 c_2$, where $c = c_1 c_2, c_1, c_2 \in \Sigma^+$. Observe that for every $c_3, c_4 \in \Sigma^*$ with $c = c_3 c_4$, there are two possibilities: if $|c_3| \leqslant |c_1|$ (with $|c_2| \leqslant |c_4|$), then there are $c_5, c_6 \in \Sigma^*$ (with $c_5 = c_3$) having $c_1 c_2 = c_5 c_4, c_1 = c_3 c_6$, and thus $(c_1 c_2 c_1 =)\ a^m = c_5 c_4 c_3 c_6$. If $|c_4| < |c_2|$ (with $|c_1| < |c_3|$), then there are $c_5, c_6 \in \Sigma^*$ (with $c_6 = c_4$) having $c_2 = c_5 c_4, c_1 c_2 = c_3 c_6$, and thus $(c_2 c_1 c_2 =)\ b^n = c_5 c_4 c_3 c_6$. Clearly, then $|a|, |b| < |c|$. Therefore, applying Proposition 11, $c_4 c_3$ is bordered. Using Theorem 2.6, $(c =) c_3 c_4 \in Q$ implies $c_4 c_3 \in Q$. Hence, because of $c \in Q$, for a suitable pair c_3, c_4 with $c = c_3 c_4$, it holds that $c_4 c_3$ is a Lyndon word. Then, by Proposition 10, $c_4 c_3$ is unbordered, a contradiction.

Case 2: $k = 2$, with $m, n \geqslant 2$, where $a^m b^n = c^k, a, b, c \in Q$ is assumed as before.

Let $c \in Q$ be a word with a minimal length satisfying this equality for some $a, b \in Q$. If $|a^m| = |b^n|$, then $a^m = b^n = c$ contradicting $c \in Q$. Therefore, we may suppose $|a^m| \neq |b^n|$.

Let, say, $|a^m| > |b^n|$. Then $a^m = cc_1, b^n = c_2$ for some pair c_1, c_2 of nonempty words with $c = c_1 c_2$. Thus, $c_1^2 b^n = c_1 c$, obviously. Therefore, using $c_1^2 b^n = c_1 c$ and $cc_1 = a^m$, by Lemma 7 we obtain $(qp)^m = c_1^2 b^n$ for some $p, q \in \Sigma^*$ with $a = pq, qp \in Q$.

If $m \geqslant 3$, then we have already proved before that this equality implies $qp = b = \sqrt{c_1}$. Using $c_1^2 b^n = c_1 c$, this leads to $(qp)^{2\ell+n} = (qp)^\ell c$ for some $\ell \geqslant 1$, i.e. $c = (qp)^{\ell+n}$ contrary to $c \in Q$. Therefore, $m = 2$ should hold.

Then $(qp)^2 = c_1^2 b^n$ so that $|qp| < |c|$. This contradicts the assumption that c is a word with a minimal length having $a^2 b^n = c^2$ for some $a, b \in Q$.

Suppose $|a^m| < |b^n|$. Then $a^m = c_1, b^n = c_2 c$ for some pair c_1, c_2 of nonempty words with $c = c_1 c_2$. Thus, $a^m c_2^2 = cc_2$, obviously. Therefore, using $b^n = c_2 c$ and $a^m c_2^2 = cc_2$, by Lemma 7 we obtain $(qp)^n = a^m c_2^2$ for some $p, q \in \Sigma^*$ with $b = pq, qp \in Q$, which leads to contradictions as before. □

Acknowledgements

This work was supported by grants of Japanese Society for Promotion of Science (Nos. L04710 and P04028), Xerox Foundation UAC (No. 1478-2004), USA, and the Hungarian National Foundation for Scientific Research (OTKA T049409). The authors are grateful to Professor Masami Ito for his constant support during their research visit in Kyoto. The authors are also grateful to the anonymous referees for their helpful comments.

References

[1] J.-P. Allouche, J. Shallit, Automatic Sequences. Theory, Applications, Generalizations, Cambridge University Press, Cambridge, 2003.
[2] D.D. Chu, T. Hsiang-Sheng, Another proof on a theorem of Lyndon and Schützenberger in a free monoid, Schoochow J. Math. 4 (1978) 143–146.
[4] P. Dömösi, S. Horváth, M. Ito, On the connection between formal languages and primitive words, in: Proc. First Session on Scientific Communication, Univ. of Oradea, Analele Univ. of Oradea, Fasc. Mat., Oradea, Romania, 1991, pp. 59–67.
[5] P. Dömösi, S. Horváth, M. Ito, Formal languages and primitive words, Publ. Math., Debrecen 20 (1993) 315–321.
[6] N.J. Fine, H.S. Wilf, Uniqueness theorems for periodic functions, Proc. Amer. Math. Soc. 16 (1965) 109–114.
[7] T. Harju, D. Nowotka, The equation $x^i = y^j z^k$ in a free semigroup, Semigroup Forum 68 (2004) 488–490.
[8] T. Harju, D. Nowotka, On unique factorizations of primitive words, TUCS Rep. 714 (2005) ⟨http://www.tucs.fi/publications/⟩.
[9] M. Lothaire, Combinatorics on Words, Addison-Wesley, Reading, MA, 1983.

[10] R.C. Lyndon, M.P. Schützenberger, The equation $a^M = b^N c^P$ in a free group, Michigan Math. J. 9 (1962) 289–298.
[11] J. Manuch, Defect theorems and infinite words, Ph.D. Dissertation, Turku, Finland, 2002.
[12] H.J. Shyr, Free Monoids and Languages, Ho Min Book Company, Taiwan, 1991.
[13] H.J. Shyr, G. Thierrin, Disjunctive languages and codes, FCT'77, Lecture Notes in Computer Science, Vol. 56, Springer, Berlin, 1977, pp. 171–176.