The Equation $a^M = b^N c^P$ in a Free Semigroup

Tero Harju
Dirk Nowotka
Turku Centre for Computer Science, TUCS,
Department of Mathematics, University of Turku

Turku Centre for Computer Science
TUCS Technical Report No 561
October 2003
ISBN 952-12-1246-2
ISSN 1239-1891

Abstract

The equation $a^M = b^N c^P$ has only periodic solutions in a free semigroup. This result was first proven by Lyndon and Schützenberger. We present a very short proof of this classical result. Moreover, we establish that the power of two or more of a primitive word cannot be factorized into conjugates of a different word.

Keywords: combinatorics on words, free semigroup, word equations

TUCS Laboratory
Discrete Mathematics for Information Technology

The equation $a^M = b^N c^P$ has only periodic solutions in a free semigroup. This result was first proven by Lyndon and Schützenberger [5] for free groups which implies the case for free semigroups since every free semigroup can be embedded in a free group. This classical result received a lot of attention. In particular, direct proofs of the subcase for free semigroups of Lyndon and Schützenbergers result were proposed; see for example [3, 2, 6]. We give a very short proof of this subcase here.

Let A be a finite set and A^* be the free monoid generated by A where ε denotes the identity. Let $A^+ = A^* \setminus \{\varepsilon\}$. Let $\mathbb{N}$ denote the set of natural numbers. We call A *alphabet*, elements of A^* *words*, the identity ε the *empty word*. We will use a, b, c, d, e, f and g to denote words, x and y for letters, and M, N, P, R, S and T for natural numbers, in the following. Let a^* denote the set of all finite powers of a, and let $a^+ = a^* \setminus \{\varepsilon\}$. Let $|a|$ denote the *length* of a. Note, that $|\varepsilon| = 0$. A word a is called *primitive*, if $a = b^M$ implies that $M \leq 1$. A word a is called *bordered*, if there exists nonempty words b, c, and d such that $a = bc = db$, otherwise a is called *unbordered*. We call two words a and b *conjugates*, if $a = cd$ and $b = dc$ for some c and d. Let $|a|_b$ denote the number of occurences of b in a. An occurrence of a in b^M is a *b-cover* of a, if $|b^{M-1}| < |a| \leq |b^M|$. Note, that a has a b-cover such that $|b| < |a|$, if, and only if, a is bordered.

Let us recall some well-known facts. We have for primitive words a that $a = bc = cb$ implies $b = \varepsilon$ or $c = \varepsilon$. It follows that for two primitive words b and c, with $|b| < |c|$, that c can have at most one b-cover. Let a be primitive. Then $bc = a^M$ implies that $cb = d^M$ where d is primitive and a conjugate of a. Note, that any primitive word a has an unbordered conjugate and that every conjugate of a occurs in aa, take for example the Lyndon word.

We now come to the announced proof.

Theorem 1 (Lyndon & Schützenberger). *If $a^M = b^N c^P$ holds with $M, N, P \geq 2$, then there exists a word d such that $a, b, c \in d^*$.*

Proof. Assume without restriction of generality that a, b, and c are primitive. The case is clear, if $b^Q \in a^+$ or $c^R \in a^+$ for some $1 \leq Q \leq N$ and $1 \leq R \leq P$. Suppose there is no $d \in A^*$ such that $a, b, c \in d^*$. So, let $b^Q \notin a^+$ and $c^R \notin a^+$ for any $1 \leq Q \leq N$ or $1 \leq R \leq P$.

If $|b| > |a|$ or $|c| > |a|$, then b or c has more than one a-cover, respectively, and hence, is not primitive; a contradiction. So, let $|b| < |a|$ and $|c| < |a|$.

If $M > 2$ then $a^M = a_0 f^{M-1} a_1$ where f is an unbordered conjugate of $a = a_0 a_1$. But, f is a factor of b^N or c^P, and thus bordered; a contradiction.

If $M = 2$ then we can assume, by symmetry, that $|b^N| > |c^P|$. Assume also that $|a|$ is of minimal length. Now, $b^N = ae^S$ and $e^S c^P = a$ for some primitive word e. From $e^{2S} c^P = e^S a$ and $b^N = ae^S$ follows that $e^{2S} c^P = g^N$

for some primitive word g. We have that $N = 2$ by the previous paragraph. But now, $|g| = |b| < |a|$ contradicts the minimality of a which proves the claim. $\square$

Lyndon and Schützenberger's result can be generalized in several ways. For example, Lentin [4] considered the equation $a^M = b^N c^P d^Q$, and Appel and Djorup [1] investigated $a^M = b_1^M b_2^M \cdots b_N^M$, where we shall mention that the solutions of these equations are not necessarily periodic, that is, their variables are not powers of the same word like in the case of $a^M = b^N c^P$.

Note, that the case $M > 2$ of the proof of Theorem 1 gives an immediate proof for the following fact for a more general equation.

Proposition 2. *Let $N \geq 2$ and $a, b_P \in A^*$, for all $1 \leq P \leq N$, be primitive. If $a^M = b_1^{M_1} b_2^{M_2} \cdots b_N^{M_N}$ with $M, M_P \geq 2$, for all $1 \leq P \leq N$, then for every $1 \leq P \leq N$ either $|b_P^{M_P}| < |a| + |b_P|$ or b_P and a are conjugates.*

Moreover, Theorem 3 below shows that not every b_P in the above proposition can be a conjugate of a. We prove that the power of a primitive word cannot be factorized into conjugates of a primitive word.

Theorem 3. *Let $N \geq 2$ and $a, b_P \in A^*$, for all $1 \leq P \leq N$, be primitive. If $a^M = b_1 b_2 \cdots b_N$ with $M \geq 2$ and b_P is a conjugate of a primitive word b, for all $1 \leq P \leq N$, then $N = M$ and $b_P = a$, for all $1 \leq P \leq N$.*

Proof. Assume that the claim does not hold, and consider a shortest counter example a for which $N \neq M$ in the statement of the theorem (over some alphabet A). By primitivity of b we can assume that $\gcd(M, N) = 1$. Indeed, if $Q = \gcd(M, N)$ then $b_1 b_2 \cdots b_{\frac{N}{Q}} = a^{\frac{M}{Q}}$ and we have equivalent solutions to the original equation. Now, $|a|/N = |b|/M \in \mathbb{N}$, and thus we can write $a = a_1 a_2 \cdots a_N$, where $|a_P| = |a|/N$ for each $1 \leq P \leq N$, and $b_R \in \{a_1, a_2, \ldots, a_N\}^M$, for all $1 \leq R \leq N$. The minimality assumption on a yields that each factor a_P is a letter. Let then $x \in A$ be a letter that occurs S times in a with $1 \leq S \leq N$, and thus MS times in a^M. Since $|b_1|_x = |b_T|_x$ whenever $1 \leq T \leq N$, we have that N divides MS, and hence, N divides S, i.e., $N = S$. But now $a = x^N$ which is a contradiction. $\square$

Finally, let us remark that the words in an equation of the form

$$a_1 a_2 \cdots a_N = b_1 b_2 \cdots b_M$$

where a_P is a conjugate of a primitive word a, with $1 \leq P \leq N$, and b_Q is a conjugate of a primitive word b, with $1 \leq Q \leq M$, are not necessarily all powers of the same word as the following example shows. Let $N = 2$ and $M = 3$ and

$$\begin{aligned} a_1 &= xxyyxy & \qquad b_1 &= xxyy \\ a_2 &= yxyxxy & \qquad b_2 &= xyyx \\ & & \qquad b_3 &= yxxy \end{aligned}$$

then

$$a_1 a_2 = xxyyxyyxyxxy = b_1 b_2 b_3.$$

References

[1] K. I. Appel and F. M. Djorup. On the equation $z_n^1 z_n^2 \cdots z_n^k = y^n$ in a free semigroup. *Trans. Amer. Math. Soc.*, 134:461–470, 1968.

[2] Ch. Choffrut. *Sec. 9.2*, volume 12 of *Encyclopedia of Mathematics and its Applications*, pages 164–168. Addison-Wesley, Reading, MA, 1983.

[3] D. D. Chu and H. Sh. Town. Another proof on a theorem of Lyndon and Schützenberger in a free monoid. *Soochow J. Math.*, 4:143–146, 1978.

[4] A. Lentin. Sur l'équation $a^M = b^N c^P d^Q$ dans un monoïd libre. *C. R. Acad. Sci. Paris*, 260:3242–3244, 1965.

[5] R. C. Lyndon and M. P. Schützenberger. The equation $a^M = b^N c^P$ in a free group. *Michigan Math. J.*, 9:289–298, 1962.

[6] J. Maňuch. *Defect Theorems and Infinite Words.* TUCS Dissertations, number 41, Turku Centre of Computer Science, University of Turku, Finland, 2002.

Turku Centre for Computer Science
Lemminkäisenkatu 14
FIN-20520 Turku
Finland

http://www.tucs.fi

University of Turku
- **Department of Information Technology**
- **Department of Mathematics**

Åbo Akademi University
- **Department of Computer Science**
- **Institute for Advanced Management Systems Research**

Turku School of Economics and Business Administration
- **Institute of Information Systems Science**