Rostock. Math. Kolloq. **59**, 71–73 (2005)

Subject Classification (AMS)
68Q45,68Q70

Alexander Meduna, Roman Lukáš

A Note on Iteratively Extendable Strings*

ABSTRACT. This scientific note introduces the notion of an iteratively extendable string within a language. It demonstrates that every language that has such an iteratively extendable string z contains infinitely many strings whose length is divisible by the length of z. Some consequences and applications of this result are given.

KEY WORDS. Formal languages, Pumping lemmas, Primes

1 Introduction

Consider a language, L. A string of the form $u_0v_1u_1v_2...u_{n-1}v_nu_n$ in L, where $v_1v_2...v_n$ is non-empty, is iteratively extendable within L if $u_0v_1^mu_1v_2^m...u_{n-1}v_n^mu_n$ is also in L, for every $m \geq 0$. In this scientific note, we prove that if there exists an iteratively extendable, z, within L, then L contains infinitely many strings whose length is divisible by the length of z. As a consequence of this, L contains infinitely many strings whose length differs from any prime. Thus, if there is a pumping lemma for a language family, such as the pumping lemma for the family of ETOL languages of finite index, then every infinite language in this family contains infinitely many strings whose length differs from any prime.

2 Definitions

This paper assumes that the reader is familiar with the theory of formal languages (see [1, 2, 3, 5]). For an alphabet, V, V^* represents the free monoid generated by V under the operation of concatenation. The identity of V^* is denoted by ε. Set $V^+ = V^* - \{\varepsilon\}$; algebraically, V^+ is thus the free semigroup generated by V under the operation of concatenation. For $w \in V^*$, $|w|$ denotes the length of w.

*Acknowledgement. This work was supported by GACR grant GA201/04/0441 and GACR grant GA102/05/H050.

Now, we introduce the notion of an iteratively extendable string within a language. Let $L \subseteq V^*$. A string $w \in L$ is *iteratively extendable within* L if $w = u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n$ for some $n \geq 1$, where $u_i, v_j \in V^*$, $0 \leq i \leq n$, $1 \leq j \leq n$, $|v_1 v_2 ... v_n| \geq 1$ and $u_0 v_1^m u_1 v_2^m ... u_{n-1} v_n^m u_n \in L$ for all $m \geq 0$.

3 Results

Theorem 3.1 *Let L be a language over an alphabet V. For every iteratively extendable string $z \in L$, there exists an infinite language $L_z \subseteq L$ such that for each $x \in L_z$, $|x|$ is divisible by $|z|$.*

Proof: Let L be a language over an alphabet V. Let z be an iteratively extendable string in L. That is, $z = u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n$ for some $n \geq 1$, where $u_i, v_j \in V^*$, $0 \leq i \leq n$, $1 \leq j \leq n$, $|v_1 v_2 ... v_n| \geq 1$ and $u_0 v_1^m u_1 v_2^m ... u_{n-1} v_n^m u_n \in L$ for all $m \geq 0$. Set $L_z = \{u_0 v_1^j u_1 v_2^j ... u_{n-1} v_n^j u_n : j = i.|u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n| + 1 \text{ for } i \geq 0\}$. Clearly, L_z is infinite and $L_z \subseteq L$. Consider any string $u_0 v_1^j u_1 v_2^j ... u_{n-1} v_n^j u_n \in L_z$ with $j = i.|u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n| + 1$ for some $i \geq 0$. Observe that $|u_0 v_1^j u_1 v_2^j ... u_{n-1} v_n^j u_n| = |u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n| + |v_1^{j-1}| + |v_2^{j-1}| + ... + |v_n^{j-1}| = |u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n| + (j-1).|v_1| + (j-1).|v_2| + ... + (j-1).|v_n| = |u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n| + (j-1).|v_1 v_2 ... v_n| = |u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n| + i.|u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n|.|v_1 v_2 ... v_n| = |u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n|.(1 + i.|v_1 v_2 ... v_n|) = |z|.(1 + i.|v_1 v_2 ... v_n|)$. Thus, Theorem 3.1 holds. ■

Corollary 3.2 *Let L be a language and $z \in L$ be iteratively extendable string; then, L contains infinitely many strings whose length is divisible by $|z|$.*

To demonstrate some applications of the previous corollary, recall that almost every textbook about formal languages proves that $\{a^n : n \text{ is a prime}\}$ is not regular in a rather complex way (c.f. Example 3.2 in [1], Example 8.8 in [2], Example 4.1.3 in [3], and Example 7.3.2 in [5]). Notice, however, that Corollary 3.2 immediately implies this result because every infinite regular language contains infinitely many iteratively extendable strings as follows from the regular pumping lemma (see Section 4.1 in [3]). From a broader perspective, if there is a pumping lemma for a language family, then this family contains no infinite language in which the length of every string equals a prime, such as $\{a^n : n \text{ is a prime}\}$. To illustrate, consider the pumping lemma for the family of ETOL languages of finite index (see Theorem 3.13 in [4]). That is, let $G = (V, P, S, \Sigma)$ be an ETOL system of index k (for some $k \geq 1$) and let $L(G)$ be infinite. Then, there exist positive integers e and $\overline{e}$ such that, for every string w in $L(G)$ that is longer than e, there exists a positive integer $n \leq 2k$ such that w can be written in the form $w = u_0 v_1 u_1 v_2 ... u_{n-1} v_n u_n$ with $|v_i| < \overline{e}$ for $1 \leq i \leq n$, $|v_1 v_2 ... v_n| \geq 1$ and for every positive integer m, the string $u_0 v_1^m u_1 v_2^m ... u_{n-1} v_n^m u_n \in L$. By Corollary 3.2 above, this family does not contain any infinite language in which the length of each string equals a prime.

References

[1] **Hopcroft, J. E.**, and **Ullman, J. D. :** *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley, Reading, 1979

[2] **Martin, J. C. :** *Introduction to Languages and the Theory of Computation.* McGraw-Hill, New York 1991

[3] **Meduna, A. :** *Automata and Languages: Theory and Applications.* Springer, London 2000

[4] **Rozenberg, G.**, and **Salomaa, A. :** *The Mathematical Theory of L Systems.* Academic Press, New York 1980

[5] **Sudkamp, T. A. :** *Languages and Machines.* Addison-Wesley, Reading 1988

received: March 26, 2003

Authors:

Alexander Meduna
Brno University of Technology,
Faculty of Information Technology,
Department of Information Systems,
Božetěchova 2, Brno 61266,
Czech Republic

e-mail: meduna@fit.vutbr.cz

Roman Lukáš
Brno University of Technology,
Faculty of Information Technology,
Department of Information Systems,
Božetěchova 2, Brno 61266,
Czech Republic

e-mail: lukas@fit.vutbr.cz