Results concerning E0L and C0L power series

Juha Honkala *

Abstract

By a classical result of Ehrenfeucht and Rozenberg the families of E0L and C0L languages are equal. We generalize this result for E0L and C0L power series satisfying the ε-condition which restricts the coefficients of the empty word.

1 Introduction

A celebrated result from classical theory of Lindenmayer systems states that the families of E0L languages and C0L languages are equal (see Ehrenfeucht and Rozenberg [1], Rozenberg and Salomaa [5,6]). In this paper we generalize this result for formal power series. We will work in the framework of morphically generated formal power series introduced in Honkala [2,3] and Honkala and Kuich [4].

In what follows A will always be a commutative ω-continuous semiring (see [4]). Suppose Σ is a finite alphabet. The set of *formal power series with noncommuting variables* in Σ and coefficients in A is denoted by $A \ll \Sigma^* \gg$. The subset of $A \ll \Sigma^* \gg$ consisting of all series with a finite support is denoted by $A < \Sigma^* >$. Series of $A < \Sigma^* >$ are referred to as *polynomials*. A semialgebra morphism $h : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ is specified by the polynomials $h(\sigma)$, $\sigma \in \Sigma$. If $h(\sigma)$ is quasiregular for all $\sigma \in \Sigma$, the semialgebra morphism h is called *propagating*. If Δ is a finite alphabet, a semialgebra morphism $h : A < \Sigma^* > \longrightarrow A < \Delta^* >$ is called a *coding* if for each $\sigma \in \Sigma$ there exist a nonzero $a \in A$ and a letter $x \in \Delta$ such that $h(\sigma) = ax$.

We are going to discuss 0L, P0L, E0L, EP0L and C0L power series. By definition, a power series $r \in A \ll \Sigma^* \gg$ is called a *0L power series* if there exist $a \in A$, $w \in \Sigma^*$ and a semialgebra morphism $h : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ such that

$$r = \sum_{n=0}^{\infty} a h^n(w). \tag{1}$$

If in (1) the semialgebra morphism h is propagating, r is called a *P0L power series*. E0L and EP0L power series are now defined in the natural way (see Honkala and

*Research supported by the Academy of Finland Department of Mathematics, University of Turku, FIN-20014 Turku, Finland, email: juha.honkala@utu.fi
and Turku Centre for Computer Science (TUCS), Lemminkäisenkatu 14, FIN-20520 Turku, Finland

Kuich [4]). A power series $r \in A \ll \Delta^* \gg$ is called an *E0L* (resp. *EP0L*) *power series* if there are a finite alphabet Σ and a 0L (resp. P0L) power series $s \in A \ll \Sigma^* \gg$ such that

$$r = s \odot \mathrm{char}(\Delta^*).$$

Finally, a power series $r \in A \ll \Delta^* \gg$ is called a *C0L power series* if there exist a finite alphabet Σ, a 0L power series $s \in A \ll \Sigma^* \gg$ and a coding $g : A < \Sigma^* > \longrightarrow A < \Delta^* >$ such that

$$r = g(s).$$

If $A = \mathbf{B}$ where $\mathbf{B} = \{0, 1\}$ is the Boolean semiring, $r \in \mathbf{B} \ll \Sigma^* \gg$ is a 0L (resp. P0L, E0L, EP0L, C0L) power series if and only if the support of r is a 0L (resp. P0L, E0L, EP0L, C0L) language. (Here the empty set is regarded as a 0L (resp. P0L, E0L, EP0L, C0L) language.)

In order to generalize the E0L=C0L theorem for formal power series it is useful to consider separately three parts of the result corresponding to different steps in its proof (see Rozenberg and Salomaa [5]; recall also that here two languages are regarded as equal if they contain the same nonempty words.)

Theorem 1 *Every C0L language is an E0L language.*

Theorem 2 *Every E0L language is an EP0L language.*

Theorem 3 *Every EP0L language is a C0L language.*

In the sequel we will generalize Theorems 1 and 3 for quasiregular power series over any commutative ω-continuous semiring A. To generalize Theorem 2 we have to introduce an additional condition. As a consequence we obtain a power series generalization of the E0L=C0L theorem.

2 C0L power series are E0L power series

In this section we prove a power series generalization of Theorem 1.

Theorem 4 *Suppose $r \in A \ll \Delta^* \gg$ is a quasiregular C0L power series. Then r is an E0L power series.*

Proof. Suppose

$$r = \sum_{n=0}^{\infty} agh^n(w)$$

where $h : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ is a semialgebra morphism, $g : A < \Sigma^* > \longrightarrow A < \Delta^* >$ is a coding, $a \in A$ and $w \in \Sigma^*$. Without restriction we assume that $\Sigma \cap \Delta = \emptyset$. Extend g and h to semialgebra morphisms $g, h : A < (\Sigma \cup \Delta)^* > \longrightarrow A < (\Sigma \cup \Delta)^* >$ by $g(x) = h(x) = 0$ if $x \in \Delta$. Next, choose a new letter $\$ \notin \Sigma \cup \Delta$ and define the semialgebra morphism $f : A < (\Sigma \cup \Delta \cup \$)^* > \longrightarrow A < (\Sigma \cup \Delta \cup \$)^* >$ by

$$f(x) = \$h(x) + g(x), \qquad f(\$) = \varepsilon,$$

$x \in \Sigma \cup \Delta$. We claim that there exist polynomials $r_n, p_n \in A < (\Sigma \cup \Delta \cup \$)^* >$, $n \geq 1$, such that

$$f^n(w) = r_n + gh^{n-1}(w) + p_n \tag{2}$$

and

$$\mathrm{proj}_{\Sigma\cup\Delta}(r_n) = h^n(w), \ \ p_n \odot \mathrm{char}((\Sigma \cup \$)^*) = 0, \ \ p_n \odot \mathrm{char}(\Delta^*) = 0 \tag{3}$$

if $n \geq 1$. (Here $\mathrm{proj}_{\Sigma\cup\Delta} : A < (\Sigma \cup \Delta \cup \$)^* > \longrightarrow A < (\Sigma \cup \Delta)^* >$ is the projection mapping \$ into ε and x into itself if $x \in \Sigma \cup \Delta$.) Clearly, there exist $r_1, p_1 \in A < (\Sigma \cup \Delta \cup \$)^* >$ such that (2) and (3) hold for $n = 1$. Suppose then that (2) and (3) hold for $n \geq 1$. Then

$$f^{n+1}(w) = f(r_n + gh^{n-1}(w) + p_n) = f(h^n(w)) = r_{n+1} + gh^n(w) + p_{n+1}$$

for suitable $r_{n+1}, p_{n+1} \in A < (\Sigma \cup \Delta \cup \$)^* >$ satisfying

$$\mathrm{proj}_{\Sigma\cup\Delta}(r_{n+1}) = h^{n+1}(w),$$

$$p_{n+1} \odot \mathrm{char}((\Sigma \cup \$)^*) = 0, \ \ p_{n+1} \odot \mathrm{char}(\Delta^*) = 0.$$

This concludes the proof of the existence of the polynomials r_n, p_n, $n \geq 1$.

Now, because

$$\sum_{n=0}^{\infty} af^n(w) \odot \mathrm{char}(\Delta^*) =$$

$$aw \odot \mathrm{char}(\Delta^*) + a\sum_{n=1}^{\infty}(r_n + gh^{n-1}(w) + p_n) \odot \mathrm{char}(\Delta^*) =$$

$$\sum_{n=0}^{\infty} agh^n(w) = r,$$

r is a E0L power series. □

3 E0L power series satisfying the ε-condition

In this section we generalize Theorem 2 for E0L power series satisfying the ε-condition. Suppose

$$r = \sum_{n=0}^{\infty} ag^n(w) \odot \mathrm{char}(\Delta^*)$$

is an E0L power series where $g : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ is a semialgebra morphism, $a \in A$, $w \in \Sigma^*$ and $\Delta \subseteq \Sigma$. We say that r satisfies the *ε-condition* if

$$(g(c), \varepsilon) = (g^n(c), \varepsilon)$$

for all $n \geq 1$, $c \in \Sigma$.

Theorem 5 *Suppose $r \in A \ll \Delta^* \gg$ is a quasiregular E0L power series satisfying the ε-condition. Then r is an EP0L power series.*

Proof. Suppose

$$r = \sum_{n=0}^{\infty} a g^n(w) \odot \mathrm{char}(\Delta^*)$$

where $g : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ is a semialgebra morphism, $a \in A$, $w \in \Sigma^*$ and $\Delta \subseteq \Sigma$. Define the semialgebra morphism $\beta : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ by $\beta(c) = (g(c), \varepsilon)\varepsilon$ for $c \in \Sigma$. Then we have $\beta(v) = (g(v), \varepsilon)\varepsilon$ for $v \in \Sigma^*$. Let $\overline{\Sigma} = \{\overline{c} \mid c \in \Sigma\}$ be a new alphabet. Define the mapping $\phi : A < \Sigma^* > \longrightarrow A < (\Sigma \cup \overline{\Sigma})^* >$ by

$$\phi(\varepsilon) = 0,$$

$$\phi(c_1 \ldots c_m) = c_1 \ldots c_m + [\beta(c_1) + \overline{c}_1] \ldots [\beta(c_m) + \overline{c}_m] - \overline{c}_1 \ldots \overline{c}_m - \beta(c_1 \ldots c_m)$$

if $m \geq 1$ and $c_1, \ldots, c_m \in \Sigma$, and

$$\phi(P) = \sum (P, w)\phi(w)$$

if $P \in A < \Sigma^* >$. (Here A is not a ring but the meaning of the subtraction above should be clear.) Next, define the propagating semialgebra morphism $h : A < (\Sigma \cup \overline{\Sigma})^* > \longrightarrow A < (\Sigma \cup \overline{\Sigma})^* >$ by

$$h(c) = h(\overline{c}) = \phi(g(c))$$

for $c \in \Sigma$. Finally, define the semialgebra morphism $\pi : A < (\Sigma \cup \overline{\Sigma})^* > \longrightarrow A < \Delta^* >$ by $\pi(c) = c$ if $c \in \Delta$ and $\pi(c) = 0$ if $c \notin \Delta$.

Now, we claim that

$$\pi h^n(c) + \beta(c) = \pi g^n(c), \tag{4}$$

$$\pi h^n(\overline{c}) + \beta(c) = \pi g^n(c) \tag{5}$$

and

$$\pi h^n(\phi(v)) + \beta(v) = \pi g^n(v) \tag{6}$$

for $c \in \Sigma$, $v \in \Sigma^+$ and $n \geq 1$. First, it is easy to see that (4) and (5) hold if $n = 1$. Suppose (4) and (5) hold for $n \geq 1$. Let $v = c_1 \ldots c_m$ where $m \geq 1$ and $c_1, \ldots, c_m \in \Sigma$. Then

$$\pi h^n(\phi(v)) + \beta(v) = \pi h^n(c_1 \ldots c_m) + \pi[\beta(c_1) + h^n(\overline{c}_1)] \ldots$$
$$\pi[\beta(c_m) + h^n(\overline{c}_m)] - \pi h^n(\overline{c}_1 \ldots \overline{c}_m) = \pi g^n(c_1 \ldots c_m).$$

Next, we have

$$g(c) = \beta(c) + \sum_{u \neq \varepsilon} (g(c), u)u.$$

Because $\beta(c) = (g(c), \varepsilon)\varepsilon = (g^2(c), \varepsilon)\varepsilon = \beta(g(c))$, we obtain

$$\beta(c) = \beta(c) + \sum_{u \neq \varepsilon} (g(c), u)\beta(u).$$

Hence

$$\pi h^{n+1}(c) + \beta(c) = \pi h^n(h(c)) + \beta(c) + \sum_{u\neq\varepsilon}(g(c),u)\beta(u) =$$

$$\pi h^n(\sum_{u\neq\varepsilon}(g(c),u)\phi(u)) + \beta(c) + \sum_{u\neq\varepsilon}(g(c),u)\beta(u) =$$

$$\pi g^n(\sum_{u\neq\varepsilon}(g(c),u)u) + \beta(c) = \pi g^n(g(c)) = \pi g^{n+1}(c).$$

Therefore (4) holds if n is replaced by $n+1$. A similar argument shows that (5) holds if n is replaced by $n+1$. This proves (4),(5) and (6) for all $n \geq 1$.

Let now \$ be a new letter and extend h and π by $h(\$) = \phi(w)$, $\pi(\$) = 0$. Then the extended h is propagating and

$$r = \sum_{n=0}^{\infty} a\pi g^n(w) = a\pi(w) + \sum_{n=1}^{\infty} a\pi h^n(\phi(w)) = \sum_{n=0}^{\infty} a\pi h^n(\$),$$

where we have used the fact that $a\beta(w) = 0$. Hence r is an EP0L power series. □

4 EP0L power series are C0L power series

To generalize Theorem 3 we need two lemmas.

Lemma 1 *If $a \in A$ and $w \in \Sigma^*$ is a nonempty word, the monomial aw is a C0L power series.*

Proof. Define the semialgebra morphism $h : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ by $h(c) = 0$ for all $c \in \Sigma$. Then

$$aw = \sum_{n=0}^{\infty} ah^n(w)$$

is a 0L power series. Hence aw is also a C0L power series. □

Note that the proof of Lemma 1, although very simple, is completely different than the language-theoretic proof that $\{w\}$ is a 0L language. In fact, the use of 0-images is unavoidable in Lemma 1. For example, if $\sigma \in \Sigma$, $\sigma \in \mathbb{N} \ll \Sigma^* \gg$ is not a 0-free C0L power series although it clearly is a 0-free EP0L power series.

Lemma 2 *If $r_1, \ldots, r_t \in A \ll \Delta^* \gg$ are quasiregular C0L power series, so is $r_1 + \ldots + r_t$.*

Proof. It suffices to consider the case $t = 2$. Let

$$r_j = \sum_{n=0}^{\infty} g_j h_j^n(a_j w_j)$$

where $h_j : A < \Sigma_j^* > \longrightarrow A < \Sigma_j^* >$ is a semialgebra morphism, $g_j : A < \Sigma_j^* > \longrightarrow A < \Delta^* >$ is a coding, $a_j \in A$ and $w_j \in \Sigma_j^*$, $j = 1, 2$. Without restriction we suppose that $a_1 \neq 0$ and $\Sigma_1 \cap \Sigma_2 = \emptyset$. Denote $k = |w_1|$ and let $\$_1,\ldots,\$_k$ be new letters. Let h be the common extension of h_1 and h_2 satisfying

$$h(\$_1) = h_1(a_1w_1) + a_2w_2, \qquad h(\$_2) = \ldots = h(\$_k) = \varepsilon.$$

Finally, let g be the common extension of g_1 and g_2 satisfying

$$g(\$_1 \ldots \$_k) = a_1g_1(w_1).$$

(The existence of g is clear if $a_1g_1(w_1) \neq 0$ or $k \neq 1$. If $a_1g_1(w_1) = 0$ and $k = 1$, we have to increase the value of k by 1.) Then

$$\sum_{n=0}^{\infty} gh^n(\$_1 \ldots \$_k) = g(\$_1 \ldots \$_k) + \sum_{n=0}^{\infty} gh^n(h_1(a_1w_1) + a_2w_2) =$$

$$a_1g_1(w_1) + \sum_{n=0}^{\infty} g_1h_1^{n+1}(a_1w_1) + \sum_{n=0}^{\infty} g_2h_2^n(a_2w_2) = r_1 + r_2$$

showing that $r_1 + r_2$ is indeed a C0L power series. □

Theorem 6 *If $r \in A \ll \Delta^* \gg$ is a quasiregular EP0L power series then r is a C0L power series.*

Proof. Suppose

$$r = \sum_{n=0}^{\infty} ah^n(w) \odot \mathrm{char}(\Delta^*)$$

where $h : A < \Sigma^* > \longrightarrow A < \Sigma^* >$ is a propagating semialgebra morphism, $a \in A$ and $w \in \Sigma^*$. Without restriction we assume that $a = 1$.

For a letter $c \in \Sigma$, the *existential spectrum* of c, denoted by espec(c), is defined by

$$\mathrm{espec}(c) = \{n \geq 0 \mid h^n(c) \odot \mathrm{char}(\Delta^*) \neq 0\}.$$

If $c \in \Sigma$, the set espec(c) is ultimately periodic, see Rozenberg and Salomaa [5,6]. (Here we use König's Lemma to avoid the difficulties caused by products equal to zero.) The threshold and period of espec(c) are denoted by thres(espec(c)) and per(espec(c)), respectively. If espec(c) is infinite, then c is called a *vital letter*. The set of vital letters of Σ is denoted by vit(Σ).

The *uniform period* associated to r is the smallest positive integer p such that
(i) for all $j \geq p$, if c is not a vital letter, then $h^j(c) \odot \mathrm{char}(\Delta^*) = 0$;
(ii) if c is a vital letter, then $p >$ thres(espec(c)) and per(espec(c)) divides p.

Let $0 \leq k < p$ and denote

$$\Sigma_k = \{c \in \Sigma \mid p + k \in \mathrm{espec}(c)\}.$$

Define the propagating semialgebra morphism $g_k : A < \Sigma_k^* > \longrightarrow A < \Sigma_k^* >$ by

$$g_k(c) = h^p(c) \odot \mathrm{char}(\Sigma_k^*),$$

$c \in \Sigma_k$. Furthermore, define the propagating semialgebra morphism $g_{p+k} : A < \Sigma_k^* > \longrightarrow A < \Delta^* >$ by

$$g_{p+k}(c) = h^{p+k}(c) \odot \mathrm{char}(\Delta^*),$$

$c \in \Sigma_k$. Note that $g_{p+k}(c) \neq 0$ for all $c \in \Sigma_k$. We claim that

$$h^{p+k}(h^p)^n h^p(P) \odot \mathrm{char}(\Delta^*) = g_{p+k} g_k^n [h^p(P) \odot \mathrm{char}(\Sigma_k^*)] \tag{7}$$

for any $n \geq 0$ and $P \in A < \Sigma^* >$. First,

$$h^{p+k} h^p(P) \odot \mathrm{char}(\Delta^*) = h^{p+k}[h^p(P) \odot \mathrm{char}(\Sigma_k^*)] \odot \mathrm{char}(\Delta^*) +$$
$$h^{p+k}[h^p(P) \odot \mathrm{char}(\Sigma^+ - \Sigma_k^*)] \odot \mathrm{char}(\Delta^*) =$$
$$h^{p+k}[h^p(P) \odot \mathrm{char}(\Sigma_k^*)] \odot \mathrm{char}(\Delta^*) = g_{p+k}[h^p(P) \odot \mathrm{char}(\Sigma_k^*)].$$

Hence (7) holds if $n = 0$. Suppose then that (7) holds for $n \geq 0$. Then

$$h^{p+k}(h^p)^{n+1} h^p(P) \odot \mathrm{char}(\Delta^*) = h^{p+k}(h^p)^n h^p[h^p(P) \odot \mathrm{char}(\Sigma_k^*)] \odot \mathrm{char}(\Delta^*) =$$
$$g_{p+k} g_k^n [h^p[h^p(P) \odot \mathrm{char}(\Sigma_k^*)] \odot \mathrm{char}(\Sigma_k^*)] = g_{p+k} g_k^{n+1}[h^p(P) \odot \mathrm{char}(\Sigma_k^*)].$$

Consequently, (7) holds for all $n \geq 0$. Therefore

$$r = \sum_{n=0}^{2p-1} h^n(w) \odot \mathrm{char}(\Delta^*) + \sum_{k=0}^{p-1} \sum_{n=0}^{\infty} h^{p+k}(h^p)^n h^p(w) \odot \mathrm{char}(\Delta^*) =$$
$$\sum_{n=0}^{2p-1} h^n(w) \odot \mathrm{char}(\Delta^*) + \sum_{k=0}^{p-1} \sum_{n=0}^{\infty} g_{p+k} g_k^n [h^p(w) \odot \mathrm{char}(\Sigma_k^*)].$$

By Lemmas 1 and 2 it suffices to prove that the series

$$s_{k,y} = \sum_{n=0}^{\infty} g_{p+k} g_k^n (y)$$

is a C0L power series if $0 \leq k < p$ and $y \in \Sigma_k^+$. For the proof fix k and y.

Next, choose nonzero polynomials P_x, $x \in \Sigma_k$, and a coding α such that

$$\alpha(P_x) = g_{p+k}(x),$$

no two of P_x, $x \in \Sigma_k$ contain a common variable, each variable of P_x has a unique occurrence in P_x and every nonzero coefficient of P_x equals 1, $x \in \Sigma_k$. Denote $P_\varepsilon = \varepsilon$ and $P_v = P_{v_1} P_{v_2} \ldots P_{v_m}$ if $m \geq 1$, $v = v_1 \ldots v_m$ and $v_i \in \Sigma_k$ for $1 \leq i \leq m$. By our choice of P_x, there exists a semialgebra morphism f such that

$$f(P_x) = \sum_{v \in \Sigma_k^*} (g_k(x), v) P_v,$$

if $x \in \Sigma_k$. Then

$$f(P_u) = \sum_{v \in \Sigma_k^*} (g_k(u), v) P_v \tag{8}$$

for any nonempty word $u \in \Sigma_k^*$. Indeed, (8) holds if $u \in \Sigma_k$ and, if (8) holds for $u \in \Sigma_k^+$ we have

$$f(P_{ux}) = f(P_u) f(P_x) =$$

$$\sum_{v_1 \in \Sigma_k^*} (g_k(u), v_1) P_{v_1} \cdot \sum_{v_2 \in \Sigma_k^*} (g_k(x), v_2) P_{v_2} = \sum_{v \in \Sigma_k^*} (g_k(ux), v) P_v$$

where $x \in \Sigma_k$.

Next, we claim that

$$f^n(P_y) = \sum_{v \in \Sigma_k^*} (g_k^n(y), v) P_v \tag{9}$$

for $n \geq 1$. First, if $n = 1$, (9) follows from (8). Suppose that (9) holds for $n \geq 1$. Then

$$f^{n+1}(P_y) = \sum_{u \in \Sigma_k^*} (g_k^n(y), u) f(P_u) =$$

$$\sum_{u \in \Sigma_k^*} (g_k^n(y), u) \sum_{v \in \Sigma_k^*} (g_k(u), v) P_v = \sum_{v \in \Sigma_k^*} (g_k^{n+1}(y), v) P_v.$$

Hence (9) holds for all $n \geq 1$. Therefore

$$\sum_{n=0}^{\infty} \alpha f^n(P_y) = g_{p+k}(y) + \sum_{n=1}^{\infty} \sum_{v \in \Sigma_k^*} (g_k^n(y), v) g_{p+k}(v) = \sum_{n=0}^{\infty} g_{p+k} g_k^n(y) = s_{k,y}.$$

This shows that $s_{k,y}$ is indeed a C0L power series. □

Now, Theorems 5 and 6 imply the following result.

Theorem 7 *If $r \in A \ll \Delta^* \gg$ is a quasiregular E0L power series satisfying the ε-condition, then r is a C0L power series.*

The necessity of the ε-condition in Theorem 7 is an open problem.

References

[1] A. Ehrenfeucht and G. Rozenberg, The equality of E0L languages and codings of 0L languages, *Intern. J. Comput. Math.* 4 (1974) 95-104.

[2] J. Honkala, On Lindenmayerian series in complete semirings, in: G. Rozenberg and A. Salomaa, eds., *Developments in Language Theory* (World Scientific, Singapore, 1994) 179-192.

[3] J. Honkala, On morphically generated formal power series, *RAIRO Theoretical Inform. and Appl.* **29** (1995) 105-127.

[4] J. Honkala and W. Kuich, On Lindenmayerian algebraic power series, *Theoret. Comput. Sci.* **183** (1997) 113-142.

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

[6] G. Rozenberg and A. Salomaa (eds.): *Handbook of Formal Languages, Vol. 1-3* (Springer, Berlin, 1997).

Received April, 2000