Theoretical Computer Science 9 (1979) 207–220

ON COMMUTATIVE DT0L SYSTEMS

Juhani KARHUMÄKI
Department of Mathematics, University of Turku, SF-20500 Turku 50, Finland

Communicated by A. Salomaa
Received June 1978

Abstract. Restricted versions of DT0L systems, so-called commutative DT0L systems, are considered. In these systems the length of a word derived is independent of the order of tables used. It turns out that many interesting length sets or languages, such as the set of composite numbers, are generated by these systems. Moreover, this approach makes it possible to give new (and slightly generalized) proofs for some undecidability results concerning DT0L functions.

1. Introduction

In recent years much attention has been paid in formal language theory to parallel rewriting. One but not the only reason for this is that the study of L systems has revealed some mathematically very interesting language families, such as the family of ET0L languages, see e.g. [7], to mention only one. The basic object behind this family is a T0L system, a parallel rewriting system with many nice properties. However, for certain purposes these systems are rather difficult to handle, mainly due to non-determinism allowed, and so the determinism restriction becomes natural. Moreover, such restricted systems, called DT0L systems, are closely related to formal power series and thus results from this latter field are available when studying DT0L (or even T0L) systems, see [9].

In this paper restricted versions of DT0L systems, so-called commutative DT0L systems, are considered. A DT0L system is called commutative if the word derived, or as in this paper its length, is independent of the order of the tables used. Especially, in a one-letter case, i.e. in a D0L case, all systems are commutative. So commutative DT0L systems may be regarded as a natural generalization of D0L systems.

The restriction put for a DT0L system to be commutative is rather strict, and so it is not surprising that all DT0L length sets are not generated commutatively. On the other hand working with commutative DT0L systems is in many respects much easier than with DT0L systems in general, and what is still more convenient is that many interesting length sets are generated, even in a natural way, commutatively by DT0L systems. For instance, we will see that for any polynomial P with nonnegative coefficients the language $\{a^{P(n_1,\ldots,n_t)} \mid n_1,\ldots,n_t \geq 0\}$ is an EDT0L language as well as the language $\{a^n \mid n \text{ is composite}\}$.

We will also show that it is decidable whether a given DT0L system is commutative (with respect to lengths of words), while some problems known to be undecidable for DT0L systems in general, see [8], remain undecidable even for polynomially bounded and commutative DT0L systems.

2. Preliminaries

In this section we define briefly the notions needed in this paper. For more detailed definitions as well as motivation the reader is referred to [3] or [7].

A *DT0L system* G is a triple $\langle \Sigma, \{h_1, \ldots, h_t\}, \omega \rangle$ where Σ is a nonempty alphabet, each h_i is a homomorphism from a free monoid generated by Σ, in symbols Σ^*, into itself, and ω is a nonempty word of Σ^*. The homomorphisms are called tables, and if $t = 1$ the system is called a *D0L system*. Let us denote

$$L_0 = \{\omega\},$$

$$L_{n+1} = \{h_1(\nu), \ldots, h_t(\nu) \mid \nu \in L_n\} \quad \text{for } n > 0.$$

Then a DT0L system G *generates the language*

$$L(G) = \bigcup_{n=0}^{\infty} L_n$$

and *the length set*

$$\text{length } L(G) = \{n \mid L(G) \text{ contains a word of length } n\}.$$

The notions like a D0L language or the family of DT0L languages are defined as usual.

Assume that $\Sigma = \{a_1, \ldots, a_r\}$. Then for a word Q in Σ^* and for a_i in Σ $|Q|_i$ denotes the number of $a_i's$ in Q and $|Q|$ denotes the length of Q. With these notions we associate with G the function F_G from $\{1, \ldots, t\}^*$ into N as follows:

For each homomorphism h_i we define an $r \times r$ matrix M_i by setting its elements $m_{qs}^{(i)}$ equal to $|h_i(a_q)|_s$. Then for each $x = i_1 \cdots i_k$ we put

$$F_G(x) = \pi M_x \eta \tag{1}$$

where π is the Parikh vector of ω, M_x denotes $M_{i_1} \cdots M_{i_k}$ and η is a column vector with all entries equal to 1. It is immediately seen that $F_G(x)$ (resp. πM_x) gives the length (resp. the Parikh vector) of the word $h_{i_k} \cdots h_{i_1}(\omega)$. So the range of F_G coincides with length $L(G)$.

Functions defined like F_G above are called in general *DT0L functions* and in the case $t = 1$ *D0L functions*. Furthermore, a DT0L function is said to be *commutative* if $F_G(x) = F_G(y)$ for all x and y with a common Parikh vector. If in (1) η is allowed, as π and $M_i's$, to be any vector with nonnegative entries, then (1) defines exactly the class of *N-rational functions*.

Above we defined the notion of commutativity with respect to lengths of words. Of course, this can be done with respect to Parikh vectors or even words, too.

Example 1. Let us consider a DT0L system G with $\Sigma=\{a_{11}, a_{12}, a_{21}, a_{22}\}$, $\omega = a_{11}a_{12}a_{21}a_{22}$ and the homomorphisms

$$\begin{aligned} h_1(a_{11}) &= a_{11}a_{12}, & h_2(a_{11}) &= a_{11}a_{21},\\ h_1(a_{12}) &= a_{12}, & h_2(a_{12}) &= a_{12}a_{22},\\ h_1(a_{21}) &= a_{21}a_{22}, & h_2(a_{21}) &= a_{21},\\ h_1(a_{22}) &= a_{22}, & h_2(a_{22}) &= a_{22}. \end{aligned}$$

Then the matrices M_1 and M_2 are

$$M_1=\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&1&1\\0&0&0&1\end{pmatrix}, \qquad M_2=\begin{pmatrix}1&0&1&0\\0&1&0&1\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

So $M_1M_2=M_2M_1$ and G is commutative. Moreover, for all x in $\{1, 2\}^*$

$$\begin{aligned} F_G(x) &= (1,1,1,1)M_1^{|x|_1}M_2^{|x|_2}(1,1,1,1)^{\mathrm{T}}\\ &= (1,1,1,1)\begin{pmatrix}1&|x|_1&0&0\\0&1&0&0\\0&0&1&|x|_1\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&|x|_2&0\\0&1&0&|x|_2\\0&0&1&0\\0&0&0&1\end{pmatrix}(1,1,1,1)^{\mathrm{T}}\\ &= (1, |x|_1+1, 1, |x|_1+1)(|x|_2+1, |x|_2+1, 1, 1,)^{\mathrm{T}}\\ &= (|x|_1+2)(|x|_2+2). \end{aligned}$$

3. Basic properties

In this section we establish some basic properties of commutative DT0L systems. First we observe that the number of different word lengths on the ith level, i.e. the cardinality of length L_i in our previous terminology, is bounded by a polynomial the degree of which is one smaller than the number of tables. More precisely,

$$\text{card length } L_i \leq \binom{i+t-1}{i}$$

where t is the number of tables of the system. So if the system contains only two tables, then card length L_i is at most $i+1$.

It follows that all DT0L functions are not generated commutatively, since in general card length L_i grows exponentially. But we can prove even more, namely that all DT0L length sets neither are generated commutatively, see Theorem 5. From this point of view it is interesting to note that the commutativity of a DT0L system is a decidable property.

Theorem 1. *It is decidable whether a given DT0L system is commutative.*

Proof. Let G be a DT0L system with matrices $M_1, \ldots, M_t$. For each r and s in $\{1, \ldots, t\}$ we define functions F_{rs} and F_{sr} from $\{1, \ldots, t\}^*$ into N by setting

$$F_{rs}(x) = \pi M_x M_r M_s \eta$$

and

$$F_{sr}(x) = \pi M_x M_s M_r \eta$$

where π stands for the Parikh vector of the axiom of G and $\eta = (1, \ldots, 1)^T$. Clearly, G is commutative iff $F_{rs} = F_{sr}$ for all r and s in $\{1, \ldots, t\}$. But these functions are N-rational and so their equality can be checked, see [1].

Our next result is very basic for this paper. It is well known, for instance, in connection with N-rational functions, see e.g. [9], and we present its proof only for the sake of completeness.

Theorem 2. *For any D0L functions $f_1, \ldots, f_t$ there exists a commutative DT0L system G with t tables such that*

$$F_G(x) = f_1(|x|_1) \cdots f_t(|x|_t).$$

Proof. Let $G_i = \langle \Sigma_i, h_i, \omega_i \rangle$ be a D0L system with the growth function f_i. Further let H_i and π_i denote the matrix determined by h_i and the Parikh vector of ω_i, respectively. We define a DT0L system G with the alphabet $\Sigma = \Sigma_1 \times \cdots \times \Sigma_t$, the Parikh vector of the axiom $\pi = \pi_1 \odot \cdots \odot \pi_t$ and the matrices

$$M_1 = H_1 \odot I \odot \cdots \odot I$$

$$M_2 = I \odot H_2 \odot \cdots \odot I$$

$$M_t = I \odot I \odot \cdots \odot H_t$$

where $\times$ and $\odot$ denotes the direct and the Kronecker product, respectively, and the identity matrices are of an appropriate size, i.e. the jth component in any Kronecker product is of the same order than H_j.

Recalling the identity $(A \odot B)(C \odot D) = (AC) \odot (BD)$ (if all products defined) we conclude that, for $r < s$,

$$M_r M_s = I \odot \cdots \odot H_r \odot \cdots \odot H_s \odot \cdots \odot I = M_s M_r.$$

Hence G is commutative. Moreover, we obtain

$$\begin{aligned}
F_G(x) &= \pi M_x \eta \\
&= (\pi_1 \odot \cdots \odot \pi_t)(M_1^{|x|_1} \cdots M_t^{|x|_t})(\eta_1 \odot \cdots \odot \eta_t) \\
&= (\pi_1 \odot \cdots \odot \pi_t)(H_1^{|x|_1} \odot \cdots \odot H_t^{|x|_t})(\eta_1 \odot \cdots \odot \eta_t) \\
&= (\pi_1 H_1^{|x|_1} \eta_1) \odot \cdots \odot (\pi_t H_t^{|x|_t} \eta_t) \\
&= (\pi_1 H_1^{|x|_1} \eta_1) \cdots (\pi_t H_t^{|x|_t} \eta_t) \\
&= f_1(|x|_1) \cdots f_t(|x|_t),
\end{aligned}$$

which proves the theorem.

Let us consider more detailed, from L systems point of view, how G above is constructed. First of all Σ may be identified with

$$\{a(i_1, \ldots, i_t) \mid i_j = 1, \ldots, \text{card } \Sigma_j\}.$$

The table associated with M_j simulates the homomorphism h_j and is defined as follows. For simplicity let $\Sigma_j = \{1, \ldots, \text{card} \rightarrow \Sigma_j\}$ and assume that $h_j(k) = k_1 \cdots k_s$. Then the table corresponding M_j contains productions

$$a(i_1, \ldots, k, \ldots, i_t) \rightarrow a(i_1, \ldots, k_1, \ldots, i_t) \cdots a(i_1, \ldots, k_s, \ldots, i_t)$$

for all possible values of $i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_t$. Finally, the axiom of G contains all the letters of the form $a(i_1, \ldots, i_t)$ where for each j i_j is a letter in ω_j. Moreover, in determining the axiom the multiplicities must be taken care of. An example of the above construction is our Example 1, where $G_1 = G_2 = \langle \{a_1, a_2\}, \{a_1 \rightarrow a_1 a_2, a_2 \rightarrow a_2\}, a_1 a_2 \rangle$.

As a generalization of Theorem 2 we immediately conclude

Corollary 1. *For any D0L functions f_{ij}, $i = 1, \ldots, t$, $j = 1, \ldots, N$, there exists a commutative DT0L system G with t tables such that*

$$F_G(x) = \sum_{j=1}^{N} f_{1j}(|x|_1) \cdots f_{tj}(|x|_t). \tag{2}$$

As regards the converse of Theorem 2 and Corollary 1 the following holds.

Theorem 3. *Any commutative DT0L function is of the form*

$$F_G(x) = \sum_{j=1}^{N} f_{1j}(|x|_1) \cdots f_{tj}(|x|_t), \tag{3}$$

where all the functions f_{ij} are N-rational and moreover the functions f_{tj} are even D0L functions.

Proof. Follows immediately from the fact that if A and B are square matrices of order N, then

$$AB = \sum_{j=1}^{N} A\eta_j\pi_j B$$

where η_j is the jth coordinate vector as a column vector and π_j is the same as a row vector.

It is known that (3), without further assumption concerning the functions f_{ij}, characterizes the set of commutative N-rational functions, see [9]. On the other hand it is clear that (3), with or without the above mentioned further assumption, does not characterize commutative DT0L functions, cf. Example 4. The formula (2) neither characterizes commutative DT0L functions, as is seen by the following example.

Example 2. We show that the function F defined by

$$F(x) = (|x|_1 + 1)(|x|_2 + 1) + (1 + (-1)^{|x|_1+1})$$

is a commutative DT0L function but is not of the form (2). The first part follows since $F = FG$ where G is a DT0L system with the alphabet $\{a_1, a_2, a_3, a_4, \bar{a}_1, a_\lambda\}$, the axiom a_1 and the tables

$$T_1: \begin{array}{l} a_1 \to \bar{a}_1 a_2 a_\lambda, \\ a_2 \to a_2, \\ a_3 \to a_3 a_4, \\ a_4 \to a_4, \\ \bar{a}_1 \to a_1 a_2, \\ a_\lambda \to \lambda, \end{array} \qquad T_2: \begin{array}{l} a_1 \to a_1 a_3, \\ a_2 \to a_2 a_4, \\ a_3 \to a_3, \\ a_4 \to a_4, \\ \bar{a}_1 \to \bar{a}_1 a_3, \\ a_\lambda \to a_\lambda. \end{array}$$

The second part is proved as follows. Assume that

$$F(x) = \sum_{i=1}^{N} f_i(|x|_1) g_i(|x|_2)$$

where f_i's and g_i's are D0L functions. Since $F(\lambda) = 1$, N must be 1. So

$$F(x) = f(|x|_1) g(|x|_2) = (|x|_1 + 1)(|x|_2 + 1) + a_{|x|_1}$$

where f and g are D0L functions and $a_{|x|_1} = (1 + (-1)^{|x|_1+1})$. Putting $|x|_1 = 0$ we obtain

$$g(|x|_2) = |x|_2 + 1$$

and hence

$$(|x|_2 + 1)(f(|x|_1) - |x|_1 - 1) = a_{|x|_1}.$$

But this is a contradiction since $(a_{|x|_1})$ is bounded but not zero-sequence. A more simple example would be the function defined by $F(x) = |x|_1 + |x|_2 + 1$.

Our next aim is to prove that all DT0L length sets are not generated commutatively. For this purpose we need some notions. We say that a language L (resp. a set K of natural numbers) contains a *gap of length N* if there exists an n such that length $L \cap \{n, \ldots, n+N\}$ (resp. $K \cap \{n, \ldots, n+N\}$) is empty. Moreover, we call a language L *strongly gapable* if there exists a real number q such that for any natural n length $L \cup \{i \mid i \geq qn\}$ contains a gap of length n. Intuitively this means that L must contain "large gaps" in the set of "short words". For instance any polynomial language L_k, i.e. $L_k = \{a^{P(n)} \mid n \geq 0\}$ where P is a polynomial of degree k, does not contain large enough gaps to be strongly gapable. This follows since the greatest gaps of L_k in the set $\{a^{P(n)} \mid n \leq m\}$ are of order m^{k-1} which is not fractional of m^k.

Finally, we define the *density of a language L*, in symbols dens L, to be

$$\lim_{n\to\infty} \frac{\text{card length } L \cap \{1, \ldots, n\}}{n}$$

if this limit exists. With these notions we prove

Theorem 4. *If a commutative DT0L language L is strongly gapable, then* dens $L = 0$.

Proof. Assume that a DT0L system G generates L. Define the family $\mathcal{F}$ by

$$\mathcal{F} = \{K \mid K \text{ is a range of a function of the form (2)}\}.$$

We shall show that length L is a finite union of sets in $\mathcal{F}$. To do this we recall that any N-rational function g over one variable can be decomposed to D0L functions, i.e. there exist a natural number p and D0L functions $g_1, \ldots, g_p$ such that

$$g(np+i) = g_i(n) \qquad \text{for } n \geq 0 \quad \text{and} \quad i = 1, \ldots, p.$$

Hence, by Theorem 3, length L is indeed a finite union of sets in $\mathcal{F}$, say length $L = K_1 \cup \cdots \cup K_n$.

To prove the theorem it suffices to show that dens $K_i = 0$ for any $i = 1, \ldots, n$. Let K_i be the range of $F_i = \sum_j f_j$ where each f_j is as in Theorem 2 a product of D0L functions, say $f_j(x) = g_{1j}(|x|_1) \cdots g_{tj}(|x|_t)$. Since K is strongly gapable so is K_i, too. Hence, for any $q = 1, \ldots, t$ there exists an index $j(q)$ such that $g_{qj(q)}$ grows exponentially. This means that F_i grows exponentially with respect to all variables, showing that

$$F_i(x) \geq A\alpha^{|x|}$$

for suitably chosen $\alpha > 1$ and A. So dens $K_i = 0$ since the number of different word lengths generated by a commutative DT0L system in less than $|x|$ steps is bounded by a polynomial in $|x|$.

Example 3. Consider a DT0L system G with the alphabet $\{a, b, c, \bar{a}, \bar{b}, a_0, b_0\}$, the axiom $\omega = a_0$ and the tables which are obtained by taking any combination of the productions listed below

$$\begin{aligned}
&a_0 \to b_0,\\
&b_0 \to a_0aaa,\ a_0aa\bar{a},\ a_0a\overline{aa},\ a_0\overline{aaa},\\
&a \to bb,\\
&\bar{a} \to \overline{bb},\ c,\\
&b \to aa,\\
&\bar{b} \to \overline{aa},\\
&c \to aaaa.
\end{aligned}$$

It is immediately seen that all words derived in $2n$ steps are of length 4^n while the length of words on levels $2n+1$ are between 4^n and $2\cdot 4^n - 1$, and in fact all of these values are reached by choosing suitable tables. This latter observation follows because all words on levels $2n$ are in $a_0\{a, \bar{a}\}^*$ and the number of $\bar{a}$'s assume (as is seen inductively) all values between 0 and $4^n - 1$. Hence

$$\text{length } L(G) = \{4^n + i \mid n \geq 0,\ i = 0, \ldots, 4^n - 1\}.$$

Evidently length $L(G)$ is strongly gapable having no density. So we obtain

Theorem 5. *All DT0L length sets are not generated commutatively.*

4. Applications

In this section we consider some consequences of the basic properties established above. Especially, we are interested in what kind of polynomials are DT0L functions of a commutative system. For this purpose we recall some notions from [4]. By a *monomial* we mean a polynomial of the form $An_1^{r_1} \cdots n_t^{r_t}$, and we say that a monomial $An_1^{r_1} \cdots n_t^{r_t}$ *covers* a monomial $Bn_1^{s_1} \cdots n_t^{s_t}$ iff $r_i \geq s_i$ for all i and $r_i > s_i$ for some i. Moreover we denote

$\mathcal{P} = \{P(n_1, \ldots n_t) \mid t \geq 1$; P is a polynomial with rational coefficients and nonnegative integer values; any monomial of P with a negative coefficient is covered by another one with a positive coefficient$\}$.

As is shown in [4] $\mathcal{P}$ characterizes N-rational polynomials, i.e. for any P in $\mathcal{P}$ there exists an N-rational function F such that

$$F(x) = P(|x|_1, \ldots, |x|_t),$$

and conversely.

Example 4. Functions

$$F_1(x)=(|x|_1+1)|x|_2+1$$

and

$$F_2(x)=(|x|_1+1)^2(|x|_2+1)-4(|x|_1+1)(|x|_2+1)+4(|x|_2+1)+1$$

and in $\mathcal{P}$ but not DT0L functions. This latter observation follows since

$$F_1(1^n2)=n+2>n=nF_1(1^n)$$

and

$$F_2(2^m11)=m+2>m=mF_2(2^m1).$$

By the above example $\mathcal{P}$ does not characterize commutative DT0L functions. However we prove

Theorem 6. *For any polynomial P in $\mathcal{P}$ there exist a commutative DT0L system G and a constant K such that*

$$F_G(x)=P(|x|_1,\ldots,|x|_t)$$

for all x satisfying $|x|_i\geqslant K$ for $i=1,\ldots,t$.

Proof. For the polynomials with integer coefficients the proof is an obvious modification of the proof of Lemma 3.1 in [4]. Hence the result follows from the

Claim. Let F be a commutative DT0L function of the form (2) and with the property that a constant p divides $F(x)$ for all x. Then there exist a DT0L function F' and a constant K' such that

$$F'(x)=\frac{1}{p}F(x)\quad \text{for } x \text{ satisfying } |x|_i\geqslant K' \text{ for all } i.$$

The proof of the claim is as follows. Let $F(x)=\sum_j f_{1j}(|x|_1)\cdots f_{tj}(|x|_t)$. We write for $|x|_i$'s large enough

$$\frac{1}{p}F(x)=\sum_j\frac{f_{1j}(|x|_1)-\alpha_j}{p}f_{2j}(|x|_2)\cdots f_{tj}(|x|_t)$$
$$+\frac{1}{p}\sum_j\alpha_j f_{2j}(|x|_2)\cdots f_{tj}(|x|_t) \tag{4}$$

where the sequences α_j are bounded and ultimately periodic keeping for all j, $f_{1j}(|x|_1)-\alpha_j$ positive-valued and making it divisible by p. For those j to whom $f_{1j}(|x|_1)$ is bounded α_j is the zerosequence and the corresponding summand in the first sum is transferred to the second one. Since the functions $(f_{1j}(|x|_1)-\alpha_j)/p$ are D0L functions, see [10] and [11], and the family of DT0L functions with a bounded range is closed under division by a constant, the claim follows from (4) by induction.

For polynomials with natural coefficients we formulate, as an immediate consequence of Theorem 2 and Corollary 1,

Theorem 7. *For any polynomial P with natural coefficients and t variables there exists a commutative DT0L system G such that*

$$F_G(x) = P(|x|_1 + 1, \dots, |x|_t + 1).$$

Corollary 2. *For any polynomial P with natural coefficients and positive values the language* $\{a^{P(n_1,\dots,n_t)} \mid n_1, \dots, n_t \geq 0\}$ *is an EDT0L language.*

Proof. It is sufficient to show that $\{P(n_1, \dots, n_t) \mid n_1, \dots, n_t \geq 0\}$ is a DT0L length set since the EDT0L languages are closed under homomorphisms. By Theorem 7, the set $\{P(n_1, \dots, n_t) \mid n_1, \dots, n_t \geq 1\}$ is a DT0L length set. So the fact that the family of DT0L length sets is closed under union guarantees that it is enough to prove that the sets of the form $\{P(n_1, \dots, n_t) \mid n_i = 0,\ n_1, \dots, n_{i-1}, n_{i+1}, \dots, n_t \geq 0\}$ are DT0L length sets. But these are ranges of polynomials with $t-1$ variables and so the result follows by induction.

Above results provide us with some rather interesting examples of DT0L length sets.

Example 5. By the identity

$$\{n \mid n \text{ is composite}\} = \{mk \mid m, k \geq 2\}$$

we conclude that the set of composite numbers is a DT0L length set, cf. Example 1. Remembering that the set of primes is not even an ET0L length set, see [5], we have found an explicit example of an EDT0L language over a one-letter alphabet whose complement is not even an ET0L language.

Example 6. The set

$$K = \{nm^2 \mid n \geq 1, m \geq 2\}$$

consists of exactly those numbers which are not square-free. So dens $K = 1 - 6/\pi^2$, see [2], showing that a DT0L language may possess a transcendental density. On the other hand, it is not difficult to see that the density of a D0L language always exists and is rational.

We finish this section by showing that all polynomials with integer coefficients are generated by commutative DT0L systems if only a dominant term is allowed, see [8]. Moreover, this dominant term can be chosen to be a very simple polynomial.

Theorem 8. *For any polynomial P with integer coefficients and t variables there exists a commutative DT0L system G and a constant K such that*

$$F_G(x) = (|x|+t+1)^K + P(|x|_1, \ldots, |x|_t).$$

Proof. Let P' be a polynomial satisfying $P(|x|_1, \ldots, |x|_t) = P'(|x|_1+1, \ldots, |x|_t+1)$. Since the function

$$(|x|+t+1)^K + P'(|x|_1+1, \ldots, |x|_t+1)$$

is, by Theorem 7, a DT0L function for great enough K, the theorem follows.

5. Undecidability results

Considerations presented in previous sections make it possible to give new (and slightly generalized) proofs for some known undecidability theorems, see e.g. [8, 9] or [1]. As in [9] we use Hilbert's tenth problem, and we believe, in a very natural way.

Theorem 9. *It is undecidable whether*

(i) *two polynomially bounded and commutative DT0L functions assume the same value for some word x,*

(ii) *for two polynomially bounded and commutative DT0L functions f and g $f(x) \geq g(x)$ holds for all words x.*

Proof. We reduce the problems to Hilbert's tenth problem which is known to be undecidable, see [6]. So assume that P is any polynomial with integer coefficients and t variables. By Theorem 8, the function $(|x|+t+1)^K + P(|x|_1, \ldots, |x|_t)$ is a DT0L function for some natural K, and so is the function $(|x|+t+1)^K$. Hence the decidability of (i) would imply the decidability of Hilbert's tenth problem.

To prove (ii) we only use functions $(|x|+t+1)^K + P(|x|_1, \ldots, |x|_t)^2$ and $(|x|+t+1)^K + 1$ instead of the above ones.

Next we turn to consider the decidability of some problems associated with one DT0L function only, such as the monotonicity problem. We are not able to prove that these problems, cf. [8], are undecidable for commutative DT0L systems, but we can do this for polynomially bounded systems. For this purpose we define like in [8] the function ODD from Σ^* into Σ^* by setting $\mathrm{ODD}(x) = x_1x_3 \cdots x_{2n-1}$ for all words of the form $x_1x_2 \cdots x_{2n}$ or $x_1x_2 \cdots x_{2n-1}$. In other words ODD catenates every second letter. Using this notion we show that polynomials over many variables are mergeable with a polynomially bounded dominant term.

Theorem 10. *For all polynomials P and Q with integer coefficients and t variables there exists a constant K and a DT0L system G such that*

$$F_G(x)=\begin{cases}(|\mathrm{ODD}(x)|+t+1)^K+P([\mathrm{ODD}(x)]) & \text{for } |x| \text{ even},\\ (|\mathrm{ODD}(x)|+t+1)^K+Q([\mathrm{ODD}(x)]) & \text{for } |x| \text{ odd}\end{cases}$$

where the square brackets denotes the Parikh vector of a word.

Proof. First we define a function S from N^t into N^t by $S(n_1,\ldots,n_t)=(n_1+1,\ldots,n_t+1)$, and let P' and Q' be polynomials with the properties $P=P'\circ S$ and $Q=Q'\circ S$. Moreover, let R be a polynomial with integer coefficients such that $P'+R$ and $Q'+R$ contain the same monomials and with positive coefficients. By Theorem 8, the function defined by

$$F_1(x)=(|x|+t+1)^K-R(|x|_1+1,\ldots,|x|_t+1)$$

is a DT0L function for some constant K. From this it follows, by the standard use of a double alphabet, that the function

$$F_2=F_1\circ \mathrm{ODD}$$

is a DT0L function, too. Observe that F_2 is not any more commutative.

Now for a polynomial T we denote by char T the polynomial containing exactly the monomials of T and with ones as their coefficients. Then as above

$$(\mathrm{char}(P'+R))\circ S\circ \mathrm{ODD} \quad \text{or equivalently} \quad (\mathrm{char}(Q'+R))\circ S\circ \mathrm{ODD}$$

is a DT0L function. But a system generating this function is easily converted, by using an extra symbol and the erasing production for it to the system generating the function F_3 defined by

$$F_3(x)=\begin{cases}((P'+R)\circ S\circ \mathrm{ODD})(x) & \text{for } |x| \text{ even},\\ ((Q'+R)\circ S\circ \mathrm{ODD})(x) & \text{for } |x| \text{ odd}.\end{cases}$$

Now the theorem follows from the identity

$$F_G=F_2+F_3.$$

As a consequence of Theorem 10 we conclude

Theorem 11. *It is undecidable*

(i) *whether a given polynomially bounded DT0L function F remains constant somewhere, i.e. whether there exist a word x and a letter b such that* $F(xb)=F(x)$;

(ii) *whether a given polynomially bounded DT0L function is monotonous.*

Proof. Quite the same as the proof of Theorem 9. Now Theorem 10 is used instead of Theorem 8.

It is interesting to note that the problems in Theorems 9 and 11 are decidable for polynomially bounded D0L systems while the status of the decidability for D0L systems in general is open.

As regards to the number of tables needed to make the above problems undecidable the following is worth mentioning. If we give up the commutativity, also in Theorem 9, then the number of tables is two. This is seen as follows. Let G be a DT0L system with more than two, say for instance with five, tables. Then there exists a DT0L system G' with two tables only such that it simulates one-step derivations of G in many (here three) steps as shown in Fig. 1, where the equality means that the corresponding words are the same. The formal definition of G' is not difficult, it is carried out by using multiple alphabets. As a consequence of this simulation it is clear that Theorems 9 and 11 are indeed valid for polynomially bounded DT0L systems with two tables only.

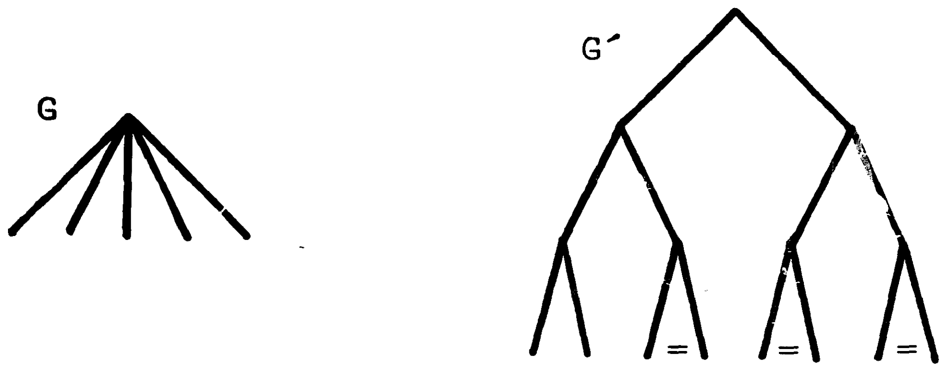

Fig. 1.

Acknowledgment

The author is grateful to Professor Arto Salomaa for useful comments.

References

[1] S. Eilenberg, *Automata, Languages and Machines* (Academic Press, New York, 1974).
[2] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers* (Oxford University Press, London, 1954).
[3] G. Herman and G. Rozenberg, *Developmental Systems and Languages* (North-Holland, Amsterdam, 1975).
[4] J. Karhumäki, Remarks on commutative N-rational series, *Theoret. Comput. Sci.* **5** (1977) 211–217.
[5] J. van Leeuwen, A study of complexity in hyper-algebraic families, in: A. Lindenmayer and G. Rozenberg, Eds., *Automata, Languages, Development* (North-Holland, Amsterdam, 1976) 323–333.
[6] Yu. Matijasevič, Enumerable sets are diophantine, *Dokl. Akad. Nauk SSSR* **191** (1970) (in Russian), *Soviet Math. Dokl.* **11** (1970) (English translation).

[7] G. Rozenberg and A. Salomaa, *The Mathematical Theory of L Systems*, to appear.
[8] A. Salomaa, Undecidability problems concerning growth in informationless Lindenmayer systems, *Elektron. Informationsverarbeit. Kybernetik* **12** (1976).
[9] A. Salomaa and M. Soittola, *Automata-Theoretic Aspects of Formal Power Series* (Springer, Berlin, 1978).
[10] M Soittola, Positive rational sequences, *Theoret. Comput. Sci.* **2** (1976) 317–322.
[11] M. Soittola, Remarks on D0L growth sequences, *RAIRO Informat. Théor.* **10** (1976).