stichting

mathematisch

centrum

ΣMC

G.T. HERMAN AND P.M.B. VITANYI

GROWTH FUNCTIONS ASSOCIATED WITH BIOLOGICAL DEVELOPMENT

Prepublication

2e boerhaavestraat 49 amsterdam

GROWTH FUNCTIONS ASSOCIATED WITH BIOLOGICAL DEVELOPMENT

by

Gabor T. Herman and Paul M.B. Vitányi.

ABSTRACT

   When an organism is growing under optimal conditions it may be assumed
that its growth rate, and that of its parts, is governed by internal, inher-
ited factors. The growth function of the organism is a function f such that
f(t) is the number of cells in the organism at time t. In the last few years
such growth functions have been actively studied by some researchers inter-
ested in mathematical models for biological development. We report on some
of the results obtained.

1. Introduction.

There have been many attempts to describe the process of biological development by mathematical models. In this article we shall only deal with mathematical models of development which are based on the approach first advocated by Lindenmayer [10, 11, 12]. Such models are commonly referred to as L systems.

The underlying idea of L systems is that the natural basic unit for the discussion of biological development is the cell. For the sake of simplicity, let us concentrate on those organisms which consist of a simple linear array of cells, referred to as a filament. (For example, many

algae are of this type. As we shall show, the mathe-
matical theory is applicable to some more complicated
organisms as well.) If we use different symbols to describe
the different states a cell may be in, then a string of
such symbols can be used to describe the whole filament.
What happens to a cell at any given time depends on its
own state and the state of its neighbors at that time. An
L system contains a set of production rules which describe
precisely how a cell changes depending on its own state
and the state of its neighbors. This change may be a
simple change of state, but it may also be a division of
the cell into two or more cells or the cell might even
disappear alltogether, i.e. die. If a string of
symbols (referred to as a word) describes the states of the
cells in a filament, then by simultaneous application of the
production rules to all the symbols in the word we obtain a
new word which describes the next stage in the development
of the filament. Repeating this process we can get the
whole developmental history of the organism. Formal
mathematical definitions of these concepts will be given in
the next section.

As will be readily noticed, the approach taken to
model development is by discretizing space and time. This
is natural in the context of biological development: we

discretize space in discrete cells and time in discrete time observations. The justification for assuming a finite set of states is that there are usually threshold values for parameters that determine the behavior of a cell. Thus, with respect to each of these parameters, it is sufficient to specify two conditions of the cell: "below threshold" and "above threshold", although the parameter itself may have infinitely many values. Even in those cases where such a simpled minded scheme is insufficient, it is usually possible to approximate the infinite set of values by a sufficiently large finite set of values, without any serious detriment to the accuracy of the developmental model.

L systems have received a great deal of attention in recent years, both because they are biologically relevant and because they are a rich source of fascinating mathematical problems. We refer the interested reader to the book of Herman and Rozenberg [7] (for a shorter account see [4]), which also contains a detailed discussion, written by Lindenmayer, of the biological significance of L systems. For recent developments in the field see Rozenberg and Salomaa [15].

In this article we shall deal with only three types of L systems: D0L systems, D1L systems and D2L systems. In all three cases the D refers to the fact that the systems are <u>deterministic</u>: in any given situation there is only one production rule which is applicable to each cell of the filament. (Although non-deterministic L systems have also been studied,

the restriction to the deterministic case is reasonable from the biological point of view, especially when we are interested in growth functions. We shall return to this point below.) The 2 in D2L systems refers to 2 sided interaction, i.e. how a cell changes depends on the states of both of its neighbors and its own state. Similarly, the 1 in D1L systems refers to 1 sided interaction: how a cell changes depends only on its own state and that of its neighbor on one side, which for a particular D1L system is always either to the left or to the right. The 0 in D0L systems refers to 0 sided interactions (or no interactions): the change of a cell is determined solely by the state of the cell itself.

An L system of any of the three types under consideration will have three components. (i) A finite nonempty set of symbols, referred to as the alphabet, which contains a symbol for all the possible cellular states between which we wish to distinguish. (ii) A set of production rules which associates with every symbol (in the D0L case), with every pair of symbols (in the D1L case) or with every triple of symbols (in the D2L case), in the alphabet a unique string of symbols by which the symbol in question will have to be replaced. (iii) A word (string of symbols) over the alphabet, referred to as the axiom, which describes the organism at the beginning of the developmental process.

If we denote the axiom by w_0 then, by applying the

production rules simultaneously to all symbols of w_0, we

obtain a string w_1. We can repeat this process any number

of times, obtaining the <u>developmental sequence</u> $w_0, w_1, w_2, w_3, \ldots$

associated with the L system under consideration. The

<u>growth function</u> associated with this L system is a function

from the natural numbers into the natural numbers, whose

value for any number t is the length of the word w_t, i.e.

the number of cells in the organism at that time.

The study of the change in size and weight of a growing

organism as a function of time constitutes a considerable

part of the literature on developmental biology. Usually,

genetically identical specimens of a specific organism are

investigated in controlled environments and their changes

of size and weight in time are described. The scientific

presupposition is that identical genetic material and identical

environment will result in identical growth rates, i.e. that

the experiment is repeatable. This assumes a deterministic

(causal) underlying structure, and makes a good case for the

biological relevance of the study of growth functions of

deterministic L systems, where we assume that the production

rules reflect the simultaneous influence of the inherited

genetic factors and a specific environment on the develop-

mental behavior of the cells. Thus, when an organism is

growing under optimal conditions it may be assumed that its

growth rate, and that of its parts, is governed by internal, inherited factors. One of the easiest things to observe about a filamentous organism is the number of cells it has. Suppose, having observed the development of a particular organism, we generalize our observations by giving a function f, such that f(t) is the number of cells in the organism after t steps. The problem then arises to produce a developmental system whose growth function is f.

This and related problems are the subject matter of this article. We conclude this section with a discussion of the biological motivation for one of the problems which will be considered in detail below. Clearly, any growth function which can be achieved by a DOL system can also be achieved by a D1L system, simply by giving production rules for the D1L system which for all practical purposes ignore the state of the neighbor. The question arises whether the converse is also true. We shall show that it is not: if a DOL system keeps growing at all, it must be growing "fast", as opposed to systems with interactions which are capable of "slow" but nevertheless unbounded growth. Thus interaction between cells provides organisms with the capability of controlling the rate of their growth in an orderly manner. When this interaction mechanism breaks down, tumors containing cells which do not interact with

their neighbors may begin to grow at an exponential rate.
For this reason, some early workers in the field of growth
functions referred to such an exponential growth as
"malignant".

2. Definitions and problem statements.

In the following Σ always denotes a finite nonempty
set of symbols, Σ^* denotes the set of all words over Σ,
including the empty word Λ (the word with no symbols in it).
If w is a word in Σ^*, then $\lg(w)$ denotes the length of
w, i.e. the number of symbols in it. In particular $\lg(\Lambda) = 0$.

Definition 1. A DxL system, $x \in \{0, 1, 2\}$, is a
construct $G = \langle \Sigma, \delta, w \rangle$, where the alphabet Σ is a
finite nonempty set of symbols; the set of production rules

δ is a mapping from $\bigcup\limits_{i=1}^{x+1} \Sigma^i$ into Σ^*, i.e. for each

ordered set $(a_1, \ldots, a_i)$ of i elements of Σ, $1 \leq i \leq x+1$,
there is one and only one α in Σ^* such that
$\delta(a_1, \ldots, a_i) = \alpha$; and the axiom w is an element of Σ^*.

Given a DxL system $G = \langle \Sigma, \delta, w \rangle$, δ induces a
mapping $\bar{\delta}$ from Σ^* into Σ^* defined as follows. $\bar{\delta}(\Lambda) = \Lambda$.

For any word $a_1 a_2 \ldots a_n \in \Sigma^*$, $n \geq 1$, $\bar{\delta}(a_1 a_2 \ldots a_n) = \alpha_1 \alpha_2 \ldots \alpha_n$ if and only if the following holds.

(i)   If $x = 0$, then $\alpha_i = \delta(a_i)$ for all i, $1 \leq i \leq n$.

(ii)   If $x = 1$, then $\alpha_i = \delta(a_{i-1}, a_i)$ for all i, $1 < i \leq n$, and $\alpha_1 = \delta(a_1)$.

(iii)   If $x = 2$, then $\alpha_i = \delta(a_{i-1}, a_i, a_{i+1})$ for all i, $1 < i < n$, $\alpha_1 = \delta(a_1)$ if $n = 1$, and $\alpha_1 = \delta(a_1, a_2)$ and $\alpha_n = \delta(a_{n-1}, a_n)$ if $n > 1$.

Hence, if $x \geq 1$, then the end cells of a filament, sensing that they have no neighboring cells on one or both sides, follow special rules. We have defined D1L systems such that each cell is influenced by its left neighbor. The case where each cell of a D1L system is influenced by its right neighbor is entirely symmetric and yields exactly the same results with respect to growth functions.

Since for an element a of Σ, $\bar{\delta}(a) = \delta(a)$, we shall from now on use the notation δ for the mapping $\bar{\delta}$ as well. Confusion is avoided by the format of the arguments.

For any natural number i, we define the i-fold composition δ^i of δ inductively by $\delta^0(v) = v$ and $\delta^{i+1}(v) = \delta(\delta^i(v))$, for each word $v \in \Sigma^*$.

Example 1. Let $G = <\{a, b, o, r\}, \delta, ar>$ be a

D1L system where δ consists of the following productions.

$\delta(o) = a$,

$\delta(x, o) = a$, for $x \varepsilon \{a, b, o, r\}$,

$\delta(a) = o$,

$\delta(o, a) = b$,

$\delta(o, b) = o$,

$\delta(o, r) = ar$,

$\delta(x, y) = y$ and $\delta(y) = y$, otherwise.

Thus,

$\delta^0(ar) = ar$,

$\delta^1(ar) = or$,

$\delta^2(ar) = aar$,

$\delta^3(ar) = oar$,

$\delta^4(ar) = abr$,

$\delta^5(ar) = obr$,

$\delta^6(ar) = aor$,

$\delta^7(ar) = oaar$, etc.

Example 2. Let $G = <\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, (,)\}$,
$\delta, 4>$ be a DOL system, where $\delta(0) = 10$, $\delta(1) = 32$, $\delta(2) = 3(4)$,
$\delta(3) = 3$, $\delta(4) = 56$, $\delta(5) = 37$, $\delta(6) = 58$, $\delta(7) = 3(9)$, $\delta(8) = 50$,
$\delta(9) = 39$, $\delta(() = ($, $\delta()) =)$.

These rules were devised in [5] in an attempt

to model without cellular interactions the developmental

behavior of certain red algae. In order to see whether we

were successful, we would have to follow the developmental sequence through quite a few steps. Because this organism grows very fast, this would be quite tedious to do by hand. We have therefore used a computer to work out the developmental sequence for us. The program CELIA (CEllular Linear Iterative Array simulator, described in [1, 2, 6]) was provided with the description of G and produced for us the developmental sequence of G, the first 16 stages of which are reproduced below.

```
1    4
2    56
3    3758
4    33(9)3750
5    33(39)33(9)3710
6    33(339)33(39)33(9)3210
7    33(3339)33(339)33(39)33(4)3210
8    33(3339)33(3339)33(339)33(56)33(4)3210
9    33(33339)33(3339)33(3339)33(3758)33(56)33(4)3210
10   33(333339)33(33339)33(33339)33(3339)33(9)3750)33(3758)33(56)33(4)3210
11   33(3333339)33(33333)33(33339)33(3339)33(39)33(9)3710)33(39)33(9)3210
12   33(33333339)33(333339)33(33333339)33(3339)33(9)3750)33(56)33(4)3210
     8)33(56)33(4)3210
13   33(333333339)33(333333339)33(33333339)33(3339)33(39)33(4)3210)33(339)33(39)33(9)3210)3
     3(33(39)33(9)3710)33(33(39)33(39)33(9)3750)33(56)33(4)3210
14   33(3333333339)33(3333333339)33(33(33339)33(33(3339)33(339)33(56)33(4)3210)33(33(3339)33(33
     39)33(39)33(9)3210)33(33(339)33(9)3710)33(33(39)33(3758)33(56)33(4)3210
15   33(3333333339)33(3333333339)33(33(3333339)33(33(33(33339)33(3339)33(3758)33(56)33(4)3210)3
     3(33(33339)33(3339)33(56)33(4)3210)33(33(3339)33(339)33(4)3210)33(339)33(39)33(9)3210)33(33(9)321
     0)33(33(39)33(9)3710)33(33(375)33(56)33(4)3210
16   33(33333333339)33(33(33333339)33(33(33(33333339)33(33(33(339)33(9)3750)33(375
     8)33(56)33(4)3210)33(33(33339)33(33(33339)33(3758)33(339)33(4)3210)33(33(33339)33(333339)3
     3(56)33(4)3210)33(33(3339)33(339)33(39)33(9)3210)33(33(339)33(39)33(9)3710)33(33(
     9)3750)33(3758)33(56)33(4)3210
```

This example is particularly interesting because it also demonstrates our earlier claim that the mathematical formalism developed can be used to investigate structures more complicated than simple linear arrays. Interpreting the left parenthesis as the beginning of a branch and right parenthesis as the end of a branch (thus parentheses within parentheses indicate branches on branches), the computer has displayed on the screen the following, which represents the first 16 stages of the development.
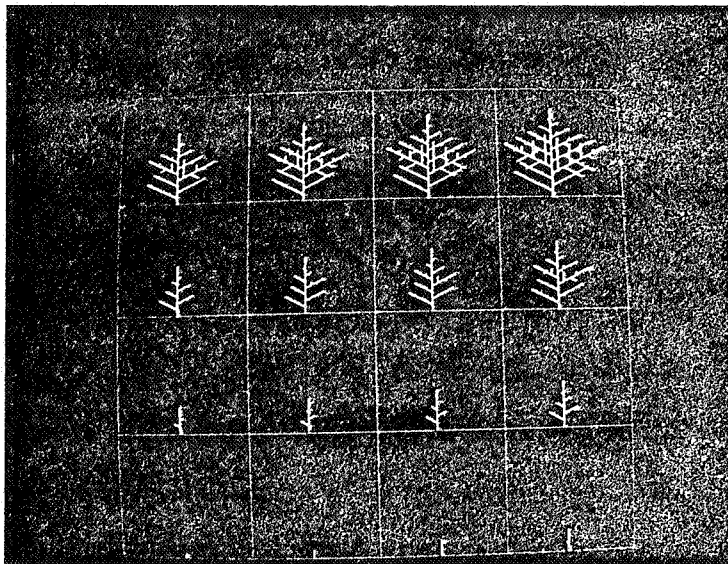
Since we wanted to have a look at the details of this development, we requested the computer to display stage 12 in some detail. This is shown below.
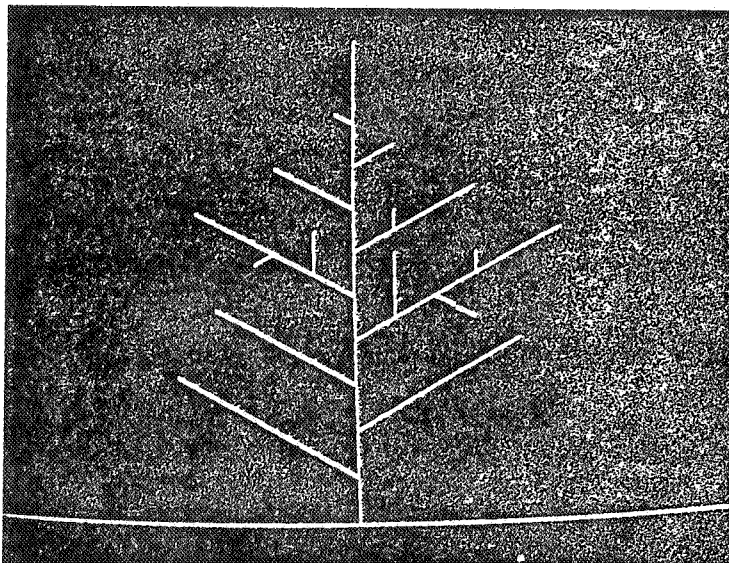
Such studies made us conclude that the development of branching patterns of certain red algae can be achieved without cellular interactions.

Examples 1 and 2 emphasize our comments at the end of the last section. In Example 1, a slow rate of growth is controlled by the 1-sided interaction, while in Example 2 the lack of interaction causes a fast rate of growth.

Definition 2. If $G = \langle \Sigma, \delta, w \rangle$ is a DxL system, $x \in \{0, 1, 2\}$, then the function f_G from the nonnegative integers into the nonnegative integers defined by

$$f_G(t) = \lg(\delta^t(w))$$

for all t, is said to be the growth function of G.

Example 3. Let $G = \langle \{a, b\}, \{\delta(a) = b, \delta(b) = ab\}, a \rangle$ be a DOL system. Then,

$$f_G(0) = f_G(1) = 1,$$

and for all t such that $t \geq 0$,

$$f_G(t + 2) = f_G(t + 1) + f_G(t).$$

Thus, $f_G(t)$ is the t'th element of the well known

Fibonacci sequence 1, 1, 2, 3, 5, 8, 13, 21, ...

Example 4. Let $G = <\{a, b, c\}, \{\delta(a) = abc^2, \delta(b) = bc^2,$

$\delta(c) = c\}, a>$ be a DOL system. Then,

$$f_G(0) = lg(a) = 1,$$

$$f_G(1) = lg(abc^2) = 4,$$

$$f_G(2) = lg(abc^2bc^4) = 9,$$

$$f_G(3) = lg(abc^2bc^4bc^6) = 16.$$

In fact, for all $t > 0$,

$$f_G(t) = f_G(t - 1) + 2t + 1.$$

By induction it follows that $f_G(t) = (t + 1)^2$.

In investigating growth functions, one of the first

questions we ask is what rates of growth are possible. That

the rate of growth of a DxL system is at most exponential

follows from the next lemma which is immediate from the

definitions.

Lemma 1. For any DxL system $G = <\Sigma, \delta, w>$, $x \in \{0, 1, 2\}$,

$$f_G(t) \leq lg(w)m^t,$$

where m is the maximum length of a value δ may have.

(I.e. $m = max\{lg(\alpha) \mid \alpha$ is in the range of the set of

production rules $\delta\}$).

The problems which have been investigated with respect to growth functions fall roughly into the following six categories.

(i) <u>Analysis problems</u>. Given a DxL system, describe its growth function in some fixed predetermined formalism.

(ii) <u>Synthesis problems</u>. Given a function f in some fixed predetermined formalism and an $x \varepsilon \{0, 1, 2\}$, find a DxL system whose growth function is f. Related to this is the problem: which functions can be growth functions of DxL systems?

(iii) <u>Growth equivalence problems</u>. Given two DxL systems, decide whether or not they have the same growth function.

(iv) <u>Classification problems</u>. Given a DxL system decide what is its growth type. (E.g., is there a polynomial or even a constant which bounds its growth function. Growth types will be rigorously defined in Section 5.)

(v) <u>Structural problems</u>. What properties of production rules induce what type of growth.

(vi) <u>Hierarchy problems</u>. Is the set of growth functions of DxL systems a proper subset of the set of growth functions of D(x+1)L systems and similar problems.

In the first five cases we would like to solve our problems effectively. That is, we would like to be able to

write computer programs (algorithms) which, in the case of
the analysis problem say, provide us with an explicit description of
the growth function whenever they are given the description
of the DxL system. (We shall return in a more rigorous way
to the concept of an algorithm in Section 4, where we shall
show that for some of the tasks described above there is
no algorithm which does the job.)

In this article we shall deal with only two questions
in detail: the analysis problem for DOL systems and the
nature of subpolynomial growth functions (these are unbounded
functions which nevertheless grow slower than any unbounded polynomial).
These are the topics of the next two sections. The section
after them summarizes some other known results about growth
functions.

3. Analysis of growth functions of DOL systems.

The next definition and lemma provide the essence of
much that we know about growth functions of DOL systems.

Definition 3. For a DOL system $G = <\Sigma, \delta, w>$,
with alphabet $\Sigma = \{a_1, a_2, ..., a_n\}$, we define the following
matrices. The initial vector π_G is the n dimensional
row vector such that its i'th component equals the number of
occurrences of the letter a_i in w, for $i = 1, 2, ..., n$.

The <u>final vector</u> η_G is the n dimensional column vector with all its components equal to 1. The <u>growth matrix</u> M_G is the $n \times n$ matrix whose (i, j)'th entry equals the number of occurrences of a_j in $\delta(a_i)$.

These matrices are introduced because from the point of view of growth of DOL systems the order of the letters in w and in the values of δ is immaterial. In fact, the following result has an easy inductive proof.

<u>Lemma 2.</u> If G is a DOL system and t is a natural number, then

$$f_G(t) = \pi_G M_G{}^t \eta_G.$$

<u>Example 5.</u> Consider the DOL system G of Example 4.

We have $\pi_G = (1, 0, 0)$, $\quad M_G = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, $\quad \eta_G = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Therefore

$$M_G{}^2 = \begin{pmatrix} 1 & 2 & 6 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix},$$

$$M_G{}^3 = \begin{pmatrix} 1 & 3 & 12 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{pmatrix} = 3M_G{}^2 - 3M_G + M_G{}^0,$$

where $M_G{}^0$ is the 3×3 identity matrix. From this and Lemma 2 it follows that

$$f_G(t + 3) = 3f_G(t + 2) - 3f_G(t + 1) + f_G(t),$$

which can be used to prove that $f_G(t) = (t + 1)^2$ for all t.

We now proceed to give an explicit formula for the growth functions of DOL systems, by using known facts concerning homogeneous linear difference equations with constant coefficients.

Theorem 1. For any DOL system, $G = \langle \Sigma, \delta, w \rangle$ the form of f_G is

$$f_G(t) = \sum_{i=1}^{k} p_i(t) c_i{}^t$$

where the c_i's are the distinct characteristic values of M_G, and p_i is an r_i'th degree polynomial in t, where $r_i + 1$ is the multiplicity of the characteristic value c_i of M_G, $1 \leq i \leq k$. (Therefore $\sum_{i=1}^{k} (r_i + 1) = \#\Sigma$).

The coefficients of the polynomials are determined by the first $\#\Sigma$ values of f_G.

$\underline{\text{Proof}}$. Let $q(x) = \sum_{i=0}^{n} a_i x^i$ be the polynomial

$\det(Ix - M_G)$, i.e. the characteristic polynomial of M_G.
By the Cayley-Hamilton Theorem, $q(M_G) = 0$ (where 0 denotes
the zero matrix). By Lemma 2,

$$\sum_{i=0}^{n} a_i f_G(t + i) = \sum_{i=0}^{n} a_i \pi_G M_G^{t+i} \eta_G$$

$$= \pi_G M_G^{t} \left(\sum_{i=0}^{n} a_i M_G^i \right) \eta_G$$

$$= 0.$$

Hence the growth function f_G satisfies a homogeneous
difference equation of order $n = \#\Sigma$ with coefficients
identical to those of the characteristic polynomial of M_G.
It is well known that such difference equations have a
solution of the form

$$f_G(t) = \sum_{i=1}^{k} p_i(t) c_i^{t}$$

where the c_i's are the distinct roots of $q(x) = 0$
(i.e. the distinct characteristic values of M_G), and the
p_i's are polynomials in t of degree r_i, where $r_i + 1$
is the multiplicity of c_i, $1 \leq i \leq k$. The coefficients
of the polynomials p_i are determined by $f_G(s)$, ..., $f_G(n - 1)$,
where s is the multiplicity of the zero root ($s = 0$ if
zero is not a root). Hence we see that the growth function
of a DOL system is a generalized exponential polynomial which
has positive integer values for positive integer arguments.

Example 6. Let $G = \langle \{a, b, c\}, \{\delta(a) = a^2,$
$\delta(b) = a^5 b, \delta(c) = b^3 c\}, a^m b^n c^p \rangle$ be a DOL system. The
characteristic equation $x^3 - 4x^2 + 5x - 2 = 0$ of the growth matrix

$$M_G = \begin{pmatrix} 2 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}$$

has roots $x_1 = x_2 = 1$ and $x_3 = 2$. (Note that M_G is
independent of the axiom.) Since the axiom has m occurrences
of a, n occurrences of b and p occurrences of c, we
obtain as the growth function of G:

$$f_G(t) = a_1 + a_2 t + a_3 2^t,$$

where

$$f_G(0) = a_1 + a_3 = m + n + p,$$
$$f_G(1) = a_1 + a_2 + 2a_3 = 2m + 6n + 4p,$$
$$f_G(2) = a_1 + 2a_2 + 4a_3 = 4m + 16n + 22p.$$

Consequently,

$$f_G(t) = (m + 5n + 15p)2^t - 12pt - 4n - 14p.$$

This shows immediately, that G has an exponentially
increasing growth function for all axioms not equal to Λ.

Example 7. In Example 5 we produced the homogeneous
difference equation

$$f_G(t + 3) - 3f_G(t + 2) + 3f_G(t + 1) - f_G(t) = 0.$$

The characteristic equation of this difference equation (and

of the matrix it derives from) is

$$x^3 - 3x^2 + 3x - 1 = 0.$$

The roots are $x_1 = x_2 = x_3 = 1$. Hence, solutions to this equation are all of the form

$$f_G(t) = at^2 + bt + c$$

Using $f_G(0) = 1$, $f_G(1) = 4$ and $f_G(2) = 9$, we obtain $a = 1$, $b = 2$ and $c = 1$, proving again that

$$f_G(t) = (t + 1)^2.$$

An alternative approach for solving the analysis problem, which is also of use for the growth equivalence problem and the synthesis problem, is an application of the theory of generating functions.

Definition 4. With any function f from the nonnegative integers into the nonnegative integers we associate its generating function $F(x)$ which is defined to be the formal infinite power series $\sum_{t=0}^{\infty} f(t)x^t$. We also say that $F(x)$ generates f.

The reason for such a definition is that very often the function $F(x)$ can be represented in a simple way. For example, if $f(t) = 2^t$ then $F(x) = \frac{1}{1-2x} = 1 + 2x + 4x^2 + 8x^3 + \ldots$.

The following lemmas are well known and easily proven mathematical facts. ($p(x)/q(x)$ denotes the fraction, $p(x)q(x)$ the product of the polynomials p and q.)

Lemma 3.

(i) If $p(x)$ and $q(x)$ are two polynomials with integer coefficients such that $q(0) = 1$, then $p(x)/q(x)$ uniquely determines an infinite power series with integer coefficients, i.e. $p(x)/q(x) = \sum\limits_{t=0}^{\infty} f(t)x^t$, where $f(t)$ is an integer for all t. Thus $p(x)/q(x)$ generates the function f. Furthermore, given $p(x)$ and $q(x)$, $f(t)$ is effectively computable for every nonnegative integer t.

(ii) Let $p(x)$, $q(x)$, $p'(x)$ and $q'(x)$ be polynomials with integer coefficients such that $q(0) = q'(0) = 1$, and let f and f' be functions generated by $p(x)/q(x)$ and $p'(x)/q'(x)$, respectively. Then $f(t) = f'(t)$ for all t if and only if $p(x)q'(x) = p'(x)q(x)$, for all x. Thus it is effectively decidable whether or not $p(x)/q(x)$ and $p'(x)/q'(x)$ generate the same function.

Lemma 4. Let n be any positive integer and let A be a $n \times n$ matrix whose entries are polynomials in x with integer coefficients. Let $q(x) = \det(A)$. If there exists a value of x such that $q(x) \neq 0$, then A is invertible, i.e. there exists an $n \times n$ matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$, where I denotes the $n \times n$ identity matrix. Furthermore, given A, A^{-1} can be effectively obtained, and each entry of A^{-1} will be of the form $p_{i,j}(x)/q(x)$, where $p_{i,j}(x)$ is a polynomial with integer coefficients.

These lemmas lead us to the following theorem.

Theorem 2. There is an algorithm which, for any DOL system G, effectively computes two polynomials $p(x)$ and $q(x)$ with integer coefficients where $q(0) = 1$, such that $p(x)/q(x)$ generates the growth function f_G of G.

Proof. Let G be the given DOL system and let π_G, M_G and η_G be as usual. Suppose the alphabet of G contains n elements. Let $M_G x$ be the $n \times n$ matrix obtained by multiplying each entry of M_G by the variable x. Let I denote the $n \times n$ identity matrix. Then $I - M_G x$ is an $n \times n$ matrix whose entries are polynomials with integer coefficients. Let $q(x) = \det(I - M_G x)$. Since $q(0) = 1$ we see that $I - M_G x$ is an invertible matrix. According to Lemma 4, we can effectively produce a $n \times n$ matrix $(I - M_G x)^{-1}$ whose entries are all of the form $p_{i,j}(x)/q(x)$, where $p_{i,j}(x)$ and $q(x)$ are polynomials with integer coefficients. Clearly, $\pi_G (I - M_G x)^{-1} \eta_G$ is of the form $p(x)/q(x)$ where $p(x)$ is a polynomial with integer coefficients and can be effectively computed. All we need to complete the proof of the theorem is to show that $p(x)/q(x)$ generates the growth function f_G of G.

For $1 \leq i \leq n$, $1 \leq j \leq n$, let $f_{i,j}$ be the function generated by $p_{i,j}(x)/q(x)$, i.e. $p_{i,j}(x)/q(x) = \sum_{t=0}^{\infty} f_{i,j}(t) x^t$. (That such an $f_{i,j}$ exists and is unique follows from Lemma 3(i).)

For $t \geq 0$, let F_t be the $n \times n$ matrix whose typical entry is $f_{i,j}(t)$. Then we have that

$$I = (I - M_G x)(I - M_G x)^{-1}$$

$$= (I - M_G x)\left(\sum_{t=0}^{\infty} F_t x^t\right)$$

$$= \sum_{t=0}^{\infty} (F_t x^t) - \sum_{t=0}^{\infty} (M_G F_t x^{t+1})$$

$$= \sum_{t=0}^{\infty} (F_t x^t) - \sum_{t=1}^{\infty} (M_G F_{t-1} x^t).$$

Identifying coefficients of powers of x we get that $F_0 = I$, and, for $t \geq 1$, $F_t = M_G F_{t-1}$. From this it follows that, for $t \geq 0$, $F_t = M_G^t$. Hence,

$$p(x)/q(x) = \pi_G (I - M_G x)^{-1} \eta_G$$

$$= \pi_G \left(\sum_{t=0}^{\infty} F_t x^t\right) \eta_G$$

$$= \pi_G \left(\sum_{t=0}^{\infty} M_G^t x^t\right) \eta_G$$

$$= \sum_{t=0}^{\infty} (\pi_G M_G^t \eta_G) x^t$$

$$= \sum_{t=0}^{\infty} f_G(t) x^t.$$

Thus, $p(x)/q(x)$ is the generating function of f_G.

This theorem can certainly be considered as a solution to the analysis problem for DOL systems, since given a DOL system the algorithm provides us with a description of its growth function in the form of a rational generating function.

Example 8. Consider the DOL system $G = <\{a, b, c\}, \delta, a>$,

where $\delta(a) = abc^2$, $\delta(b) = bc^2$, $\delta(c) = c$, of Examples 4, 5,

and 7. As in Example 5,

$$\pi_G = (1, 0, 0), \qquad M_G = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \eta_G = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Using Cramer's rule we see that

$$\pi_G (I - M_G x)^{-1} \eta_G = \frac{\det \begin{pmatrix} 1 & -x & -2x \\ 1 & 1-x & -2x \\ 1 & 0 & 1-x \end{pmatrix}}{\det \begin{pmatrix} 1-x & -x & -2x \\ 0 & 1-x & -2x \\ 0 & 0 & 1-x \end{pmatrix}}$$

$$= \frac{1+x}{(1-x)^3} = 1 + 4x + 9x^2 + 16x^3 + \ldots \quad .$$

4. Subpolynomial growth functions.

In this section we investigate some of the properties
of the class of growth functions which are unbounded and yet
grow slower than any unbounded polynomial.

Definition 5. A function is said to be unbounded if
$\lim_{t \to \infty} f(t) = \infty$.

<u>Definition 6.</u> The growth type of a DxL system

$G = <\Sigma, \delta, w>$ $(x \in \{0, 1, 2\})$ is said to be subpolynomial

if and only if f_G is unbounded and for every unbounded

polynomial p it is the case that $\lim_{t\to\infty} (f_G(t)/p(t)) = 0$.

<u>Example 9.</u> The growth type of the D1L system G

of Example 1 is subpolynomial. It is in fact easy to show

that f_G is unbounded and that, for all positive integers

k, there exists an integer t such that

$$f_G(t + 1) = \ldots = f_G(t + 2^k),$$

and so, for any unbounded polynomial p, $\lim_{t\to\infty} (f_G(t)/p(t)) = 0$.

The next theorem makes precise our claim that unbounded

growth functions of D0L systems must grow "fast".

<u>Theorem 3.</u> If $G = <\Sigma, \delta, w>$ is a D0L system and

m is an integer such that $f_G(m) = f_G(m + 1) = \ldots = f_G(m + n)$,

where n is the number of symbols in Σ, then for all

$t \geq 0$, $f_G(m + t) = f_G(m)$.

<u>Proof.</u> Let $q(x) = \sum_{i=0}^{n} a_i x^i$ be the characteristic

polynomial of M_G $(n = \#\Sigma$ and $a_n = 1)$. We prove that

$f_G(m + t) = f_G(m)$ for all $t \geq 0$ by induction. The cases

$0 \leq t \leq n$ follow from the condition in the theorem. Suppose

now that the result is valid for $0 \leq t \leq s$, where $s \geq n$.

We have shown in the proof of Theorem 1 that for all $k \geq 0$,

$\sum_{i=0}^{n} a_i f_G(k + i) = 0$. Letting $k = m + s - n + 1$, we get

using the induction hypothesis that

$$f_G(m + s + 1) = -\sum_{i=0}^{n-1} a_i f_G(m + s - n + 1 + i)$$

$$= -\sum_{i=0}^{n-1} a_i f_G(m + s - n + i)$$

$$= f_G(m + s).$$

This leads directly to the following hierarchy result.

Theorem 4. The set of growth functions of D0L systems is a proper subset of the set of growth functions of D1L systems.

Proof. That the set of growth functions of D0L systems is a subset of the set of growth functions of D1L systems follows by definition. That it is a proper subset follows from Theorem 3 and Example 9.

The essence of the proof above is that unbounded growth functions of D0L systems must grow with at least a certain speed, and we have found an unbounded growth function of a D1L system which grows slower than this. It is interesting to note that the same type of argument cannot be repeated to show the existence of a growth function of a D2L system which is not also a growth function of a D1L system. This is because the longest period for which an unbounded growth function of a D2L system $G = \langle \Sigma, \delta, w \rangle$ can retain the value k is clearly n^k where $n = \#\Sigma$. It is not too difficult to prove that for any integer $n > 1$ there is a D1L system $G = \langle \Sigma, \delta, w \rangle$ with $\#\Sigma = n + 2$, such that

f_G is unbounded and f_G retains the value k for at least n^k consecutive arguments. (Example 1 serves this purpose in case n = 2.) It is at present an open problem whether or not there exists a D2L system which has a growth function which is not also the growth function of a D1L system.

We complete this section with an example of a task for the execution of which there is no algorithm.

Theorem 5. There is no algorithm which for any given D1L system G decides whether or not the growth type of G is subpolynomial.

Proof. The proof of this theorem makes use of the theory of Turing machines. A Turing machine is a logical device consisting of a finite control with an attached read-write head travelling about on an infinitely expandable tape divided into squares. Each square contains one of a finite set of symbols, and according to the current state of its finite control and the symbol in the scanned tape square, the Turing machine prints a new symbol in the square under scan, moves one square to the left or to the right and enters a new state. If the Turing machine enters a special state, then it is said to halt. Consider the following task: give an algorithm which for any Turing machine decides whether or not that Turing machine halts if it is started on a blank tape. This is referred to as

the "blank tape halting problem".

It is standard in the theory of computation to
identify the intuitive concept of an "algorithm" with the
mathematically precise concept of a "Turing machine" (Church's
thesis). It is also well known (see, e.g., Minsky [13])
that there is no Turing machine which "solves" the blank tape
halting problem. However, it can be shown (for details see
Vitányi [19]) that if there was an algorithm which decides for
any given D1L system G whether or not the growth type of
G is subpolynomial, then it could be used to construct an algorithm
which solves the blank tape halting problem. Since the
latter does not exist, the former cannot exist either.

5. Summary of other results on growth functions.

(i) Analysis problems.

In section 3 we have discussed and solved the analysis
problem for D0L systems. There we have used two different
formalisms: sums of exponential functions with polynomial
coefficients, and rational generating functions. We have
described procedures for obtaining the growth function
in either of these formalisms for an arbitrary D0L system.

Theorem 5 and similar results have interesting
consequences regarding the analysis problem for D1L
and D2L systems. It is true that we may

be able to find a suitable mathematical expression for
the growth function of any given D2L system by ad hoc methods.
However, if we fix a formalism in which we want to express
the growth function in a way which clearly indicates the
growth type of the function, then there is no algorithm
which, for an arbitrary D1L system G,   gives an explicit
expression for f_G in the predetermined formalism.

(ii)   Synthesis problems.

A major result in the direction of synthesis of growth
functions for DOL systems is the following.   It can be
shown (Szilard [17]) that any positive, nondecreasing,
ultimately polynomial function is the growth function of a
DOL system.   The proof provides an algorithm which for any
such function (described in some predetermined way) produces
the required DOL system.

The method uses many results on the nature of
polynomial functions.   On the way to proving the main theorem
Szilard showed, for example, that if the generating functions
$F(x)$ and $F'(x)$ generate growth functions of DOL systems, then
so do $F(x) + F'(x)$, $1 + xF(x)$ and $F(x)/(1 - x)$.   His
proofs were effective, given the DOL systems whose growth
functions are generated by $F(x)$ and $F'(x)$,   he showed
how we can obtain the DOL systems whose growth functions are

generated by $F(x) + F'(x)$, $1 + xF(x)$ and $F(x)/(1 - x)$.
Thus, if we know how to obtain DOL systems whose growth
functions are generated by certain basic generating functions,
results such as this provide us with the ability to construct
DOL systems whose growth functions are generated by more and
more complicated generating functions put together from the
basic ones by the operations described above.

Such method can certainly be used to synthesize a
function like $(t + 1)^2$, i.e. to produce a DOL system whose
growth function is $(t + 1)^2$. In fact much more complicated
growth functions can also be synthesized.

Presently it is an open problem whether or not there
exists an algorithm such that given a function f, either
by an exponential polynomial as in Theorem 1 or by its
generating function as in Theorem 2, the algorithm decides
whether or not f is a DOL growth function. On the other
hand, there are algorithms which, given a DOL growth function
f in either of these formalisms, will produce a DOL system
whose growth function is f   (see, e.g., Paz and Salomaa [14]).

When we come to L systems with interaction, the situa-
ation is again much worse. There are no general algorithms
for synthesizing growth functions of D1L and D2L systems,
but some partial results have been obtained. For

example, Vitányi [19] proved that for each rational number

$0 < r < 1$, we can effectively find a D2L system whose growth

function is of the order of magnitude of t^r.

(iii) Growth equivalence problems.

It is a consequence of Theorem 2 and Lemma 3(ii) that

there is an algorithm which, for any two D0L systems, decides

whether or not they have the same growth function.

Exactly the opposite is the case for D1L and D2L

systems. Vitányi [19] proved that there does

not exist an algorithm which decides the growth equivalence

of two arbitrary D1L systems. However, such an algorithm

exists for the rather restricted case of those D1L systems

whose growth functions are bounded or which have an alphabet

of one letter only.

(iv) Classification problems.

In view of Lemma 1, the following gives an exhaustive

classification of growth types.

The growth of a DxL system $G = \langle \Sigma, \delta, w \rangle$ $(x \in \{0, 1, 2\})$

is said to be

(i) <u>exponential (type 3)</u> if and only if there exists

a real number $x > 1$ such that $\lim_{t \to \infty} f_G(t)/x^t > 0$;

(ii) <u>subexponential (type $2\frac{1}{2}$)</u> if and only if the

growth is not exponential and there does not exist a poly-

nomial p such that $f_G(t) \leq p(t)$ for all t;

(iii) <u>polynomial (type 2)</u> if and only if f_G is unbounded and there exist polynomials p and q such that $p(t) \leq f_G(t) \leq q(t)$ for all t;

(iv) <u>subpolynomial (type 1 $\frac{1}{2}$)</u> if and only if f_G is unbounded and for each unbounded polynomial p

$$\lim_{t \to \infty} f_G(t)/p(t) = 0;$$

(v) <u>limited (type 1)</u> if and only if there exists an integer m such that $0 < f_G(t) < m$ for all t;

(vi) <u>terminating (type 0)</u> if and only if there exists an integer t_0 such that $f_G(t) = 0$ for all $t \geq t_0$.

It is known [18] that growth types $2\frac{1}{2}$ and $1\frac{1}{2}$ cannot occur in the DOL case (cf. Theorem 3). There are DOL systems with growth types 3, 2, 1, 0. The system of Example 4 is clearly polynomial (type 2), while the system of Example 6 is exponential (type 3) for a nonempty axiom. Theorem 1 can also be used to find the growth type of a given DOL system:

we work out an expression of the form $\sum_{i=1}^{k} p_i(t)c_i^{\,t}$ as

described in Theorem 1, and read off what the growth type is from this expression.

There exist D1L systems with growth type $1\frac{1}{2}$, the system in Example 1 is such. Karhumäki [9] has given an example of a D2L system of growth type $2\frac{1}{2}$, and by a result in [19] this implies that there is also a D1L system of growth type $1\frac{1}{2}$.

Earlier on we have pointed out that we need interaction in order to make an unbounded growth rate "slow". More

formally this can now be expressed by saying that in the
DOL case if a growth function is not limited it must be
at least polynomial, however in the DlL case it may be
subpolynomial. Nevertheless, even in the D2L case, a growth
function which is not limited must grow with at least an
approximately logarithmic rate. Vitányi [19] proved
that if $G = <\Sigma, \delta, w>$ is a D2L system such that $r = \#\Sigma > 1$
and f_G is unbounded, then

$$\lim_{t \to \infty} [(\sum_{i=0}^{t} f_G(t)) / (\sum_{i=0}^{t} \lfloor \log_r ((r - 1)i + r) \rfloor)] \geq 1.$$

As far as growth classification problems for L systems
with interactions are concerned, the results are all negative.
Vitányi [19] proved that if $x \in \{1, 2\}$ and
$i \in \{0, 1, 1\frac{1}{2}, 2, 2\frac{1}{2}, 3\}$, then there is no algorithm which
decides for an arbitrary DxL system whether or not its growth
type is i (see Theorem 5).

(v) Structural problems.

The view we have taken until now is global in its
approach, i.e. we have not yet considered what properties
of the production rules cause the different types of growth.
We now give an example of a typical structural result.

Let $G = \langle \Sigma, \delta, w \rangle$ be a DOL system, and let a be a symbol in Σ. Then a is said to be <u>expanding</u> if and only if there exists a $t \geq 0$ and $\alpha_1, \alpha_2, \alpha_3$ in Σ^*, such that $\delta^t(a) = \alpha_1 a \alpha_2 a \alpha_3$, and a is said to be <u>accessible</u> if and only if there exists a $t \geq 0$ and α_1 and α_2 in Σ^*, such that $\delta^t(w) = \alpha_1 a \alpha_2$. Salomaa [16] proved that a DOL system $G = \langle \Sigma, \delta, w \rangle$ is exponential (type 3) if and only if there exists a symbol a in Σ which is both accessible and expanding.

Vitányi [18] proved similar results on the inter-relationship between the nature of production rules and the type of growth of DOL systems for the cases of type 0, 1 and 2. He has proved for example that it is possible to give four DOL systems which differ from each other only in their axioms, and which nevertheless are of type 0, 1, 2 and 3, respectively. In fact, if a DOL system $G = \langle \Sigma, \delta, w \rangle$ is of growth type 2, then there is a substring v of $\delta^{\#\Sigma}(w)$ such that $\langle \Sigma, \delta, v \rangle$ is of growth type 1.

Thus our knowledge of structural problems for DOL systems is rather exhaustive. As opposed to this, little work has been done on structural problems for D1L and D2L systems. In view of the results on the classification problems for such systems, an exhaustive set of solutions to the structural problems is impossible.

(vi) Hierarchy problems.

Theorem 4 is a typical hierarchy result. As we have
mentioned, the problem whether or not there exists a D2L
growth function f which is not the growth function of a
D1L system is still unknown.

A simple hierarchy result is the following [19]. If
G is a D2L system with a one letter alphabet, then f_G is
the growth function of a D0L system. However, there are
D1L systems with two letter alphabets whose growth functions
are not D0L growth functions.

Finally, let us point out that in Definition 1 we
have allowed a symbol to map into the empty string. It
is natural to consider the so called propagating L
systems, in which we do not allow production rules which
have Λ in their range. Although such restrictions can
drastically limit the possibilities of a system, much but
by no means all of the theory described above remains
valid even under the propagating restriction.

For example, the impossibilities of algorithmic
solutions to the analysis, growth equivalence and classi-
fication problems carry over to the propagating DxL systems
$(x \in \{1, 2\})$, except for the growth equivalence problem for
propagating D1L systems of subpolynomial growth type, which is

still open. There are propagating D1L systems using an
alphabet of two letters whose growth functions are not
D0L growth functions. For each propagating DxL system G,
x ε {0, 1, 2},

$$\lim_{t\to\infty} (f_G(t)/\lfloor \log_r((r-1)t+r)\rfloor) \geq 1,$$

where r is the cardinality of the alphabet of G, a
result which does not hold without the propagating restriction.

There exist propagating D2L systems whose growth
function is of the order of magnitude t^r for $0 < r < 1$.
It is conjectured [19] that there is no such propagating
D1L system. If this conjecture is true, it follows that
the set of growth functions of propagating D1L systems is
a proper subset of the set of growth functions of propagating
D2L systems.

6. Historical notes.

The first paper in the field of growth functions was
by Szilard [17] who treated the analysis and synthesis
problems for D0L systems with the generating function approach.
In Paz and Salomaa [14] growth functions of D0L systems are
investigated from the point of view of integral sequential
word functions and algorithms are obtained for the solution
of the analysis, synthesis and growth equivalence problems.

The difference equation method appears in Doucet [3], Paz
and Salomaa [14] and Salomaa [16] which latter paper contains
the closed form expression for DOL growth functions and a
classification of growth types of DOL systems together with
a result on the structure of DOL growth. The present classi-
fication of growth types, as well as most of the results on
the classification and structural problems, appears in Vitányi
[18]. The first example of à D1L system with subpolynomial
growth and Theorems 3 and 4 are due to Herman (see [8] or [14]).
Karhumäki [9] provided an example of a (propagating) D2L
system of growth type $2\frac{1}{2}$. The remainder of the results on
growth functions of L systems with interaction appearing
in this article are due to Vitányi [19], which contains further
results as well as some interesting conjectures. Further relevant
references can be found in [7].

7. Acknowledgements.

References.

1. R. Baker and G. T. Herman, Simulation of organisms using a developmental model, Part I: Basic description, International Journal of Bio-Medical Computing, 1972, v. 3, 201-215.

2. R. Baker and G. T. Herman, Simulation of organisms using a developmental model, Part II: The Heterocyst formation problem in blue-green algae, International Journal of Bio-Medical Computing, 1972, v. 3, 251-267.

3. P. G. Doucet, Growth of word length in DOL systems, in: Open House on Unusual Automata Theory, Aarhus University, 1972, Tech. Rept. DAIMI PB-15, Aarhus University, 1973.

4. G. T. Herman, Using formal language theory to model biological processes, in: Applied Computation Theory, (ed. R. T. Yeh), Prentice-Hall, 1974.

5. G. T. Herman, A. Lindenmayer, and G. Rozenberg, Description of developmental languages using recurrence systems, Mathematical Systems Theory, 1974, to appear.

6. G. T. Herman and W. H. Liu, The daughter of CELIA, the French flag and the firing squad, Simulation, 1973, v. 21, 33-41.

7. G. T. Herman and G. Rozenberg, Developmental Systems and Languages, North-Holland Publishing Co., 1974.

8. G. T. Herman and A. Walker, The syntactic inference problem as applied to biological systems, in: Machine Intelligence 7 (eds. B. Meltzer and D. Michie), Edinburgh University press, 1972.

9. J. Karhumäki, An example of a PD2L system with growth type $2\frac{1}{2}$, Information Processing Letters, 1974, to appear.

10. A. Lindenmayer, Mathematical models for cellular interactions in development, Part I: Filaments with one-sided inputs, Journal of Theoretical Biology, 1968, v. 18, 280-299.

11. A. Lindenmayer, Mathematical models for cellular interactions in development, Part II: Simple and branching filaments with two-sided inputs, Journal of Theoretical Biology, 1968, v. 18, 300-315.

12. A. Lindenmayer, Developmental systems without cellular interactions, their languages and grammars, Journal of Theoretical Biology, 1971, v. 30, 455-484.

13. M. Minsky, Computation - Finite and Infinite Machines, Prentice-Hall, 1967.

14. A. Paz and A. Salomaa, Integral sequential word-functions and growth equivalence of Lindenmayer systems, Information and Control, 1973, v. 23, 313-343.

15. G. Rozenberg and A. Salomaa (eds.), Topics in L Systems, Springer-Verlag, 1974.

16. A. Salomaa, On exponential growth in Lindenmayer systems, Indagationes Mathematicae, 1973, v. 35, 23-30.

17. A. Szilard, Growth functions of Lindenmayer systems, University of Western Ontario Computer Science Department Technical Report No. 4, 1971.

18. P. M. B. Vitányi, Structure of growth in Lindenmayer systems, Indagationes Mathematicae, 1973, v. 35, 247-253.

19. P. M. B. Vitányi, Growth of strings in context dependent Lindenmayer systems, in: Topics in L Systems (eds. G. Rozenberg and A. Salomaa), Springer-Verlag, 1974.