J. theor. Biol. (1975) **54**, 3–22

Developmental Algorithms for Multicellular Organisms: A Survey of L-Systems†

ARISTID LINDENMAYER

Theoretical Biology Group, University of Utrecht, Heidelberglaan 2, Utrecht, The Netherlands

(*Received* 20 *November* 1974)

Multicellular organisms are construed as arrays of symbols, each symbol standing for a cell. Their development is modelled by algorithmic rules which provide the allowed substitutions of new arrays for each symbol in the previous array. These substitutions stand for cell divisions, or cell death, or for changes of cellular states. The substitutions may depend on inputs from neighbouring cells (development with cellular interactions), or only on the state of the cell itself (development without interactions). By repeated use of a system of algorithms one obtains a developmental sequence of arrays, or a set of arrays (a developmental language). Additional control mechanisms may be applied to the arrays generated by a system, thus defining a hierarchy of language families. This article presents a survey of biologically relevant mathematical results available on these algorithmic systems.

1. Introduction

The development of plants and animals can be studied and described in many different ways and at different levels. At the present time most interest is centered on the molecular and macromolecular mechanisms of morphogenesis and differentiation as the two main aspects of development. At the molecular level one can speak of morphogenesis in the sense of aggregation of macromolecules into larger structures, such as membranes, microtubules and flagella. These aggregate structures, in turn, serve to determine the shapes of cellular organelles and of cells, as well as influencing the direction of cell divisions, cell enlargement and of cellular transport processes. Groups of cells acting in unison, form tissues or organs. In this manner, one eventually gets to growth descriptions of entire organisms at the gross morphological level.

† A shorter version of this paper was presented to the Conference on Biologically Motivated Automata Theory, held at McLean, Virginia, June 1974. This work has been supported by NATO research grant No. 574.

A complete description of a developing organism would have to specify the essential processes and structures in the proper time sequence at all levels of organization. This goal is not only unattainable but is actually undesirable since such a description would be so complex that nothing could be done with it. It seems to be much more promising to attempt to formulate a theory of development on the basis of some essential and unifying features. It seems to me that focusing at the behaviour of cells in the course of the development of multicellular organisms provides such a basis.

The following reasons may be given for supporting the choice of cells as basic units. (1) Cells are metabolically autonomous units of all higher organisms. (2) All metabolic and synthetic activity in a cell is mediated by the protein and RNA molecules which are produced according to particular portions of DNA (the genes). (3) Among all cell constituents only the DNA molecules can be reproduced faithfully, thus inheritance of cellular mechanisms can practically be effected only by the transmission from mother to daughter cells of particular sorts of DNA. (4) Each cell in an organism is a descendant of a single ancestral cell, the fertilized egg, and each cell of an organism carries the same complement of DNA. (5) Some of the genes constituting the DNA complement of a cell may at any one time be active (producing RNA and protein) or inactive. (6) The activation (de-repression) or inactivation (repression) of genes is brought about by molecules which are produced either in the same cell or which enter the cell from neighbouring cells or from the environment.

Since in higher organisms the cells appear to be the only functionally autonomous units with genetic continuity, we assume them to be the basic units in our developmental descriptions. The property of genetic continuity among cells and their descendants introduces another important aspect of development, namely that of developmental "programs" or "algorithms". As all cells of an organism in its entire life are descendants of a single cell, one might ask what "program" this cell carries that enables it to specify the course of development the organism follows. Again, this question may be posed for developmental processes at all different levels, but at the moment we wish to concentrate on the programming of cellular processes. The question is therefore: what kinds of "instructions" are necessary to bring about the timing and spacing of major cell processes, such as cell division, cell enlargement, cell death and differentiation, during the whole life of the organism.

This question does not necessarily have a single answer. This is because algorithms may be formulated in different ways and many different sets of algorithmic instructions can give rise to the same development. Finding one such set of instructions would not ensure that these instructions would

actually be responsible for the development of a particular organism. Nevertheless, it is not entirely hopeless to ask this question since even if one cannot be sure which set of instructions is responsible for a particular development, one can mostly rule out classes of such sets which cannot be responsible. Recognizing these classes can substantially help the experimentalist in looking for the actual mechanisms. For instance, if on the basis of theoretical considerations one can say that certain kinds of development cannot be programmed without cell interactions, then it would be futile to base experiments on the assumption that cells act independently in that organism.

This discussion of "algorithms" and "programs" presupposes that we can formulate them in a general and consistent manner, and that they are furthermore readily interpretable in the usual biochemical and cell-physiological terms. I think that the mathematical framework of arrays of finite automata fulfils these conditions. The concept of finite automata is general enough to provide for all kinds of timing cycles, transport mechanisms, inducing and inhibitory effects within and among cells that one might consider. The states of cells as automata are also readily interpretable in terms of presence or absence of chemical cell constituents (with respect to threshold concentrations), and/or in terms of combinations of active and inactive genes. The inputs of cells are the compounds which enter it during a certain time interval, or the membrane excitations it received. Similarly, their outputs are compounds which are leaving the cell or excitations it gives rise to. The state transition function is partly an expression of the effects of genes on each other, in the sense of repression or depression of genes by products of other genes, and partly of control effects among the RNA and enzyme molecules in the cytoplasm.

The mathematical constructs (L-systems) which we propose to use for describing development are of a discrete nature in several respects. First of all, we have discrete spatial units, the cells, which were discussed above.

Secondly, the states and inputs of cells are assumed to be discrete entities and there are to be finitely many of them. There are, of course, a definite number (few thousand) of discrete genes, each in a few copies in each cell. Thus, if in developmental regulations we would need to consider only the combinations of active genes which are present at any one time in a cell, each active gene giving rise to a compound important in some way to growth and morphogenesis, then the state (as well as the input and output) of each cell would naturally be a discrete entity. Should there be a large number of genes involved in developmental processes in an organism, the number of possible combinations of active and inactive genes could be extremely large. This does not in fact seem to be the case: in developmental processes which

have been investigated in detail only a small number of important compounds were found.

There is, however, the added complication that we cannot simply view development as a process in which a number of genes are switched on or off in different cells at different times. Clearly, the various cytoplasmic components may vary continuously in their concentrations, and there are many known control effects among these components, such as feedback and allosteric inhibition of enzymes, inhibition or stimulation of the enzyme synthetic rates at the RNA level, diffusion of metabolites, etc. The possibilities for cytoplasmic influence on the course of development could be hopelessly numerous and involved. There is again evidence, however, that most of the metabolic and synthetic processes are regulated to stay at steady state levels, and that developmental events, such as differentiation, take place when sudden changes occur from one steady state to another in a number of biochemical pathways. Although continuous cytoplasmic changes, which are important for development, cannot be ruled out, it is not too far-fetched to assume that in the majority of cases the cytoplasmic events can also be considered as discrete ones.

A third parameter which is assumed to be discrete in our systems is time. The computation of new cell states, based on the previous states and inputs, takes place at certain intervals. Since this interval can be chosen as small as desired, this assumption does not seem to be a serious limitation. It is a problem, however, that the smaller the time interval is, the larger the number of states necessary for various timing cycles. Clearly, a compromise must be found between very detailed and very gross descriptions. The recognition that most cell processes are highly stable, with occasional shifts into other areas of stability, makes a realistic developmental description without too finely spaced time steps a possibility.

2. The Developmental Model

According to the views presented above we describe the development of multicellular organisms by growing arrays of finite automata. Each configuration of states in such an array is a momentary description of the organism, and the whole life of the organism is described by a sequence of such configurations. The same state transition function applies to all of the automata in the array. A state transition may consist of a change of the state symbol of a cell into a new symbol, or it may represent cell division by allowing the substitution of more than one new symbol in the array in place of the previous symbol, or it may stand for cell death by letting the

empty string be substituted for a symbol. The transition function may be either deterministic (having only a single value for any given combination of present state and inputs) or non-deterministic (having more than one value for some combinations). If the transition function is deterministic, then one obtains a single sequence of arrays from a given initial array.

Deterministically produced sequences of arrays constitute an interesting class of abstract objects in their own right, besides being readily interpretable as life histories of organisms. These sequences may be finite or infinitely long. They have a finite length only if at some stage the entire array disappears (automata may be removed from as well as added to such arrays, so under certain conditions it may happen that all automata are removed and no further application of the transition function is possible). Infinitely long sequences may in turn consist of infinitely repeating periodic occurrences of some arrays (after an initial non-repeating portion of the sequence), or they consist of arrays which grow without bounds. Thus we may speak of terminating growth (in the case that the sequence is finite) or of limited growth (when it is ultimately periodic), or of proper growth (when the size of the arrays increases without bound). The theory of growth functions of one-dimensional deterministic sequences has been intensively investigated (see references further below), and is of great interest biologically.

Apart from being deterministic or non-deterministic, the state transition function may also be given either as a purely combinatorial construct, i.e., by specifying separately for each combination of states and inputs the next state or states; or the function may be specified by a general formula, or by an approximation to a continuous function (Lindenmayer, 1974). The simulation language CELIA (cellular linear iterative array generator) has been specifically designed by Baker & Herman (1972) and Herman & Liu (1973) to make use of composite state transition functions in programming growing cellular arrays. This language has been applied to simulate heterocyst differentiation in blue-green algae (Baker & Herman, 1972), and to the modelling of the development of large branching and differentiating plant structures such as inflorescences (Frijters & Lindenmayer, 1974).

The approach of using growing automata arrays in order to model development can in general be applied to any multicellular or multi-compartmental organism, whether consisting of filaments, sheets, or three-dimensional structures. However, the problem of how to specify the connections of newly substituted chunks inside two- or three-dimensional arrays has made the extension of these models to sheet-like or bulky organisms more difficult (work is in progress on "graph L-systems", Culik & Lindenmayer, 1974). Most of the work up to now has concerned either simple filaments or branching filaments. Although branching filaments are actually three-dimensional

objects, they can be easily coded into one-dimensional strings by various parenthesis notations (Lindenmayer, 1968, 1971), and the strings can be generated algorithmically (in the case when inputs are received by the cells, the input processing from the branches must be specially regulated).

The development of higher organisms is usually more conveniently simulated on the basis of segments consisting of thousands or millions of cells, rather than on the basis of individual cells. As long as the segments may be assumed to be autonomously programmed, their use as units in L-systems present us with no problems.

When only one-dimensional arrays are considered, the formalisms and results of formal language theory become applicable to our constructs. The set of states is then interpreted as the alphabet of a language generating system, the state transition function becomes the set of productions, and the starting array the axiom. The language which is generated by such a system is a set which consists of all the strings produced from the axiom under the simultaneous application of production rules to every symbol in the axiom, repeating this process on all the strings generated in the first step, and continuing for infinitely many steps. A developmental language is thus an unordered set consisting of all possible stages of a certain species of organisms one may find. Of course, this set may be either finite or infinite; finite if the development is terminating or limited along every possible generating sequence, and infinite otherwise. A developmental language corresponds to a museum or herbarium collection of specimens conserved in all possible stages of their development, and without knowing the order in which they were produced.

When in a one-dimensional automata array no inputs are taken into account between the cells, i.e., when the next-state function is a function only of the present state of each cell, then we call the corresponding language generating system an "OL-system" (these have also been called "zero-sided" or "informationless Lindenmayer systems"). When the automata receive inputs from m left neighbours and n right neighbours, i.e., when the next-state function is a mapping from a sequence of state symbols which is $m+n+1$ symbols long, then we call the language generating system an "$\langle m, n\rangle$ L-system". Clearly, a $\langle 0, 0\rangle$ -L-system is the same as an OL-system. "One-sided Lindenmayer systems" were originally defined as $\langle m, n\rangle$ L-systems such that $m = 1$ and $n = 0$, or $m = 0$ and $n = 1$, but now this definition has been modified in the sense that they are called "one-sided" if $m>0$ and $n = 0$, or $m = 0$ and $n>0$. Similarly, "two-sided Lindenmayer systems" were originally $\langle 1, 1\rangle$ L-systems, while now they are defined as $\langle m, n\rangle$ L-systems such that $m>0$ and $n>0$. All $\langle m, n\rangle$ L-systems, for whatever non-negative values of m and n, are called "IL-systems". The term

"L-systems" is meant to include these as well as other related systems some of which are mentioned below.

Productions of $\langle m, n\rangle$ L-systems are written in the form: $\langle a_1, \ldots, a_m, \underline{b}, c_1, \ldots, c_n\rangle \rightarrow \alpha$, where the underlined symbol b is the one for which the string α on the right-hand side has to be substituted. In the case of OL-systems, we simply write productions as: $b \rightarrow \alpha$.

The formal definitions of these and related systems, as well as many of the mathematical results concerning them are given by Salomaa (1973) and Herman & Rozenberg (1975). The latter book contains a chapter by me and three chapters by the authors which discuss some of the biological insights that these systems can provide, as well as suggestions for certain concrete applications to morphogenetic problems.

3. Development Without Cellular Interactions

The usefulness of OL-systems to biological problems consists primarily of gaining an understanding of the morphogenetic power of cell lineages. What is meant by this is the recognition of the type of structures which can arise by autonomously programmed cells, the behaviour of which is controlled only by their lineage (their ancestry). Each cell may change its state repeatedly, or divide equally or unequally, or die, and no effects may take place between them. The complexity of structures which can be generated in this way has been very surprising, and may be instructive to biologists looking for mechanisms to account for certain types of development. It is entirely possible that in many cases where an interactive mechanism is postulated by experimentalists, a non-interactive one would suffice. Cell lineages have, of course, been studied extensively by biologists for many years, and in cases where such an attempt has been successful we can be almost sure that we are dealing with non-interactive mechanisms (essentially OL-systems). We may mention recent studies on the green alga *Chaetomorpha linum* (Lück, 1974) as an example for finding OL-type development.

The most distinctive characteristics of deterministic OL-systems is a certain repetitiveness of the substrings which they produce (cf. Ehrenfeucht & Rozenberg, 1973; Rozenberg, 1974). This property is most obvious in the so-called "locally catenative" sequences generated by some of the deterministic OL-systems (Rozenberg & Lindenmayer, 1973). In this case one can find a formula which obtains for all the strings produced (except for a few initial ones) and which specifies each string as a concatenation of some

previous strings in the sequence. Structures produced be repeated concatenations of previous stages are known in biology as "compound structures", such as compound leaves composed of leaflets, compound inflorescences composed of flowers, etc. The repeated and nested occurrence of larger and larger components in such structures is due to cyclically occurring symbols in the derivations. The biological significance of these kinds of models lies in the recognition that compound structures can result from simple cycles of cellular states.

An extension of the locally catenative property is the "recurrence" property (Herman, Lindenmayer & Rozenberg, 1975). Recurrence systems are sets of formulas which specify (by concatenation rules) all of the strings of a given OL-sequence. A locally catenative system is a recurrence system with a single formula. Every OL-system, as well as every EOL-system (see above reference), has the recurrence property. These characteristics of interactionless developmental structures are more difficult to recognize in nature, and as far as I know have not yet received a biological name. Which does not mean that it might not in the future serve to characterize an important class of developmental processes, namely those in which there are interlocking cycles of cellular states.

As an example for the generating power of an OL-system we take the development of a compound leaf. It is well known that all cells in a leaf originate in the row of cells at the leaf margin. In order to account for the generation of the gross form of the leaf, the positions of lobes and notches, one needs therefore to program only the marginal row of cells. It is true that if interactions can take place among the cells in the course of development, then such interactions may occur via cells which lie not on the margin but somewhere on the blade of the leaf. But assuming morphogenetic mechanisms without interactions, the cells lying in the interior have no effect, and we can model leaf development by considering only the row of marginal cells.

The following deterministic OL-system produces a developmental sequence giving rise to a compound leaf. Let the symbols standing for cellular states be: a, b, c, d and k. Let the production rules (the state transition function) be as follows: $a \rightarrow cbc$, $b \rightarrow dad$, $c \rightarrow k$, $d \rightarrow a$, $k \rightarrow k$. Two of the production rules are such that one cell gives rise to three new cells. While this is violating the usual case of binary cell divisions, these rules were chosen because of symmetry considerations. It would be quite easy to construct a similar OL-system with only binary cell divisions (see Rozenberg & Lindenmayer, 1973), but more states would be needed in order to coordinate the right-hand and left-hand side productions. Let a be the starting symbol. Then we obtain the following developmental sequence:

a
cbc
kdadk
kacbcak
kcbckdadkcbck
kkdadkkacbcakkdadkk
kkacbcakkcbckdadkcbckkacbcakk
kkcbckdadkcbckkkdadkkacbcakkdadkkkcbckdadkcbckk

.
.
.

We can clearly observe that each array from the fourth array on has a repeating structure which can be expressed by the following formula for all $n \geqslant 4$,

$$S_n = kS_{n-3}S_{n-2}S_{n-3}k.$$

This is a locally catenative formula except for the appearance of the constant k at the ends of each array. Strictly speaking, this sequence of arrays is defined by two recurrence formulas, one similar to the above one, except for having the term k_{n-1} for each k, and the other formula simply stating that $K_n = K_{n-1}$, for all $n \geqslant 2$.

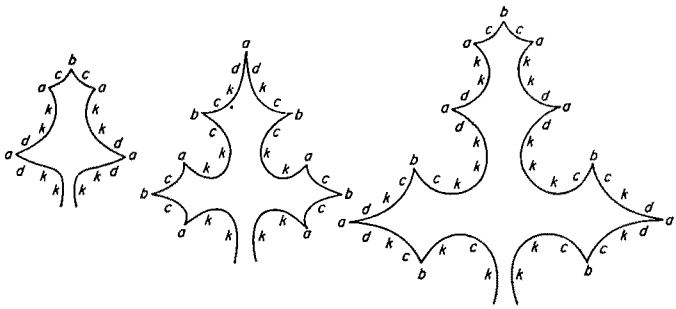

FIG. 1. Compound development of leaf-margins, generated by a locally catenative OL-system.

Our interpretation of leaf shapes represented by the last three arrays are shown in Fig. 1. We represent the symbols *a* and *b* by sharp projecting tips, the symbols *c* and *d* by the lateral margins of lobes, and the symbols *k* by notches, the older the notch the more *k*'s. The formula which describes this sequence indicates that the two side lobes repeat the whole structure of the leaf 3 steps back, and the centre lobe repeats the whole leaf 2 steps back.

This developmental behaviour is due to two cycles of states we have postulated in our production rules, namely, one cycle of length 3 going through states *a*, *b*, *d*, and another cycle of length 2 with states *a* and *b*. These cycles are shown in the state transition diagram for our system (Fig. 2).

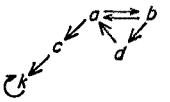

FIG. 2. State transition diagram for the locally catenative OL-system.

Thus we have demonstrated that a compound leaf can be generated by a rather simple mechanism involving five cellular states, provided that a proper cyclic state transition function is found. The generating algorithm given here gives rise to an infinite sequence of arrays, thus it programs for increasingly larger and more compound leaf shapes. Since normal leaf growth terminates at some definite stage, one would have to have a mechanism superimposed on the one given in order to stop development at some stage. Below we will mention some of the methods by which such stopping control can be imposed.

4. Development With Cellular Interactions

The systems with interactions among the cells (IL-systems) apply to a much wider range of developmental processes. Various developmental control mechanisms which have been proposed in the literature and which are based on transport or diffusion of active substances, or on propagation of excited states (e.g. of membranes) can be expressed in terms of IL-systems (provided that they act on one-dimensional structures). The insights that one can gain by translating the usual formalisms (differential equations) for such mechanisms into our discrete notation, depend on the useful mathematical results that one can obtain for IL systems. At the present time there are only a few such results. One might name here the results of Rozenberg (in press) concerning the proper hierarchy of $\langle m, n\rangle$ L-languages with increasing values of m and n. Also, his finding that the set of $\langle m, n\rangle$ L-languages is identical to the set of $\langle 1, m+n-1\rangle$ L-languages. The values of m and n needed to simulate a particular developmental process has clearly to do with the speed at which the active substances (or excitations) involved may travel in one or the other direction along the filament, or has to do with their diffusion constants.

A $\langle 1, 1\rangle$ L-system has been constructed (by Lindenmayer in Herman & Rozenberg, 1975) for the well-studied development of the main root of

maize. Erickson & Sax (1956) have obtained data concerning the rate of expansion of thin segments along the root as a function of the distance of each segment from the root tip. They could thus obtain the change of rate of growth (the relative elemental rate of growth) as a function of distance from tip. This relative elemental rate of growth is low near the tip, it increases to a rate of 40% expansion per hour at 4 mm from the tip, and then falls to zero at 10 mm from the tip.

In order to model this essentially filamentous growth behaviour, we assume interactions to take place both in the apical and in the basal directions. Time steps are taken at one hour intervals, segments of 2 mm width are taken as the basic spatial units, and their relative elemental growth rates are expressed as digitized doubling rates. For instance, segments exhibiting 40% per hour expansion rate are assumed to double twice every 5 hr. Altogether five sorts of 2 mm segments are assumed, designated from tip to base as *a*, *b*, *c*, *d*, and *e*. The timing cycles controlling doubling correspond to relative elemental rates of 1/6, 2/5, 1/4, 1/10 and 0 per hour, and are as shown in Fig. 3.

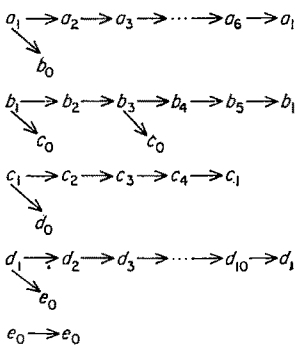

FIG. 3. The timing cycles controlling doubling.

Thus in this particular model one needs $6+5+4+10+1 = 26$ states. Changes of states are programmed by the following rules (where p and q stand for any pair of symbols from the set $\{a,b,c,d,e\}$, $p \neq q$, and $x>0$, $y \geqslant 0$).

$$\langle p_y, \underline{q_x} \rangle \to q_{x+1}$$
$$\langle \underline{q_0}, q_x \rangle \to q_{x+1}$$
$$\langle p_0, \underline{p_x}, q_y \rangle \to q_0$$
$$\langle \underline{p_0}, p_0 \rangle \to p_0$$

In each rule the change of state is given for a segment which is in a state corresponding to the underlined term as influenced by either its immediate

left or its right neighbour segment or both. These rules represent some very simple assumptions concerning physiological effects root segments may have on each other, such as the production of hormones or nutrients in one part of the root and their transmission and utilization in another part.

A developmental sequence determined by the above rules is shown in Fig. 4.

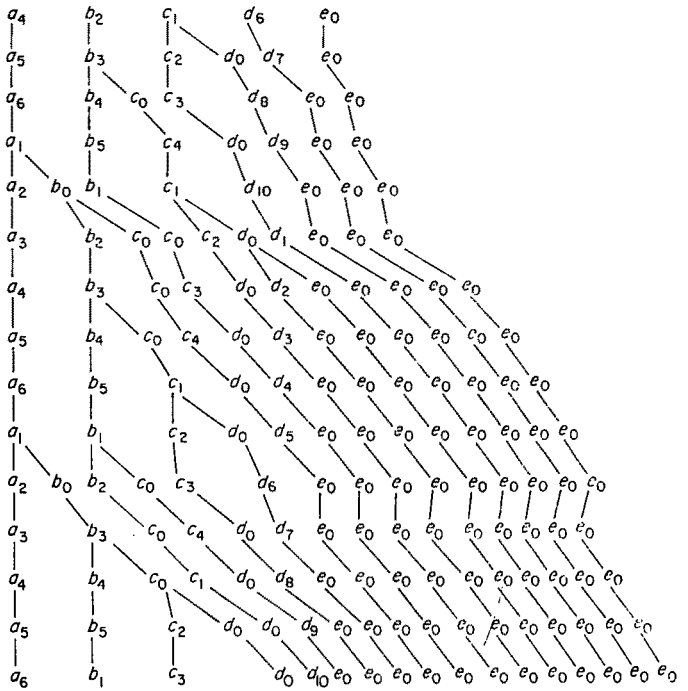

FIG. 4. Developmental sequence for the primary root of maize, generated by a ⟨1, 1⟩ L-system.

Other applications of IL-systems have been published by Baker & Herman (1972) for heterocyst distributions in the blue-green alga *Anabaena*, by Herman & Liu (1973) for shell patterns of *Conus*, and by Herman (1971, 1972) in connection with the roles of polarity in regenerating systems.

5. Environmental Effects on Development

An extension of OL-systems has been proposed by Rozenberg (1973), called "table OL-systems". In this case not only a single set of productions

may be provided in a given system, but several distinct sets of productions. These sets of productions are termed "tables", and the rule is adopted that in the generation of any given string only productions from a single table may be used. The order in which the various tables are applied to consecutive strings is not specified, but derivations may be defined in which given table sequences are required. Biologically speaking, tables of productions can represent developmental responses to changing environmental conditions.

The most important environmental effects on development are either those of a periodically changing factor, such as light or temperature or moisture, or those brought about by a single change of environment, such as photoperiodic induction of physiological processes produced when daylength exceeds or gets below certain critical value.

For a case of development under periodically changing environment a two-table OL-system given by Surapipith & Lindenmayer (1969) might be mentioned. The development of sexual reproductive structures (perithecia) in the fungus *Sordaria* could be shown to be under the influence of white light in the presence of a certain photosensitizing agent. When the fungus is growing linearly in growth tubes under alternating periods of darkness and light, successive segments of the mycelium receive different sequences of dark and light, and this results in the appearance of perithecia in bands along the growth tube. For each regime of dark and light periods one gets a particular banded pattern. Our model could account for these different patterns on the basis of a few simple state transitions.

To illustrate an application of table OL-systems to development under a single change of environment, we consider leaf margin development again, but now with the possibility of switching from one set of production rules to another. This kind of leaf development is more realistic, because now we can obtain terminating growth for each particular leaf, while the set of all leaves produced by the system is infinitely large.

We take a deterministic two-table OL-system. Let the set of state symbols consist of letters a, b, c, d, k, j, m, the starting symbol be a, and the tables of productions be:

V: $a \to kbk$, $b \to cdc$, $c \to e$, $d \to kek$, $e \to jej$, $j \to j$, $k \to k$, $m \to m$.

F: $a \to kmk$, $b \to e$, $c \to j$, $d \to kmk$, $e \to jmj$, $j \to j$, $k \to k$, $m \to m$.

Beginning with symbol a, we can derive terminal arrays corresponding to fully grown leaves by allowing development under table V for an arbitrary number of steps, and then switching to table F. Physiologically such a change could result from local unavailability of nutrients, or from photoperiodic induction to flowering condition. We show below a sample derivation.

V	*a*
V	*kbk*
V	*kcdck*
V	*kekekek*
V	*kjejkjejkjejk*
F	*kjjejjkjjejjkjjejjk*
F	*kjjjmjjjkjjjmjjjkjjjmjjjk*
	same

If we interpret, as before, the *k*'s as notches in leaf margins, the *j*'s as lateral portions or serrations of lobes, and the *m*'s as tips of lobes, then we have derived a three-lobed leaf shape with three lateral segments (*j*) on both sides of each lobe. In fact, we see that depending on the time when the table switch is done (at or after the third step), three-lobed leaves are obtained with all possible numbers of lateral segments for each lobe (but the same number of lateral segments within one particular leaf). Thus we get leaf shapes such as those shown in the class L_4 in Fig. 6.

If the switching of tables from *V* to *F* takes place before the third computation step, we get transitional shapes with one lobe only. Similarly, one lobed leaves are obtained by sequences of tables beginning with *VFV*. Some of these leaf forms are shown in Fig. 5.

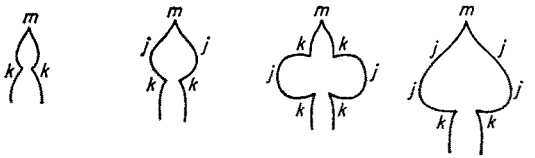

FIG. 5. Transitional forms of leaf-margins, generated by a table OL-system.

These kinds of transitional forms of leaves are well known in plants which have undergone photoperiodic induction, producing a number of reduced leaves before the apex is converted into flower parts. Similar table-switching mechanisms may be responsible for plants with heteroblastic sequences (a shoot producing leaves of different shapes in the course of a growing season).

The adult language produced by the above deterministic table OL-system under any sequences of tables is the following:

$$\{kjkmkjk\} \bigcup \{kj^n mj^n k | n \geqslant 0\} \bigcup \{kj^n mj^n kj^n mj^n kj^n mj^n k | n \geqslant 1\}.$$

Of course, many other sets of leaf shapes, with simple or multiple lobes and terminating growth can be produced analogously by table-OL-systems.

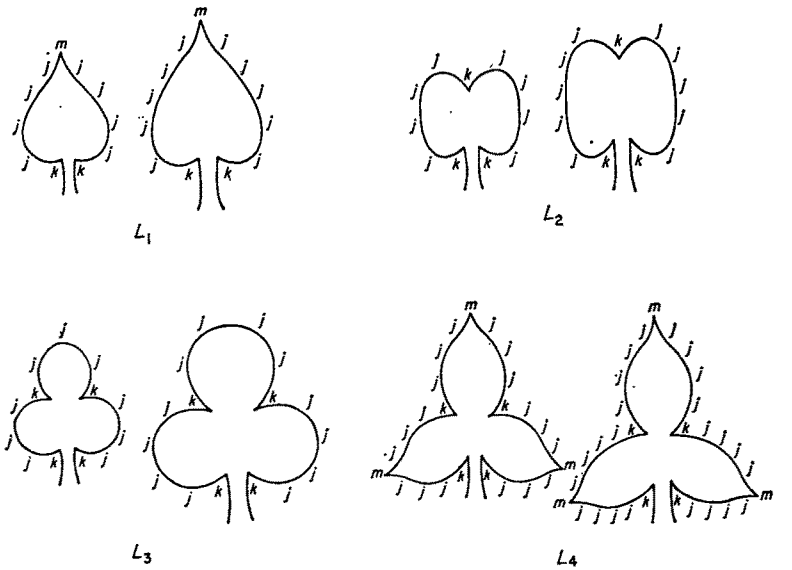

FIG. 6. Sets of leaf-shapes: L_1, symmetric entire leaves with tips; L_2, leaves with two rounded lobes of same size; L_3, leaves with three rounded lobes of same size; L_4, leaves with three lobes with tips of the same size.

The kinds of shapes with terminating growth which cannot be produced by OL-systems without tables will be discussed below.

Table L-systems with interactions have been employed to model the timing and placing of flowers in inflorescence development (Frijters & Lindenmayer, 1974). The role of environment in IL-systems has been explored by Herman (1970).

6. Development with an Adult Stage

We have seen in the previous section that adult stages of development can be represented by arrays which consist of cells in "terminal" states, i.e., states which give rise only to themselves. But it is also possible to have terminating development when the states of cells change, nevertheless the array as a whole remains unchanged. For instance, some cells may die at each step and their place is taken by others produced at that step. This occurs normally in epithelial tissues, for instance.

One can then select from a particular L-language those arrays which derive themselves and only themselves. Such a set of arrays is called the "adult language" of an L-system (Herman & Walker, in press).

Some very interesting results were obtained by Herman & Walker from

the biological point of view. Namely, the set of adult languages of OL-systems is exactly identical to the family of context-free languages (sets of strings generated by a context-free Chomsky grammar). Furthermore, the set of adult languages of right-linear OL-systems (those OL-systems which have productions only of the form $a \rightarrow c$ or $a \rightarrow bc$ where c is either a terminal symbol or the empty symbol, and a and b are non-terminal symbols) is exactly identical to the family of regular languages (sets of strings generated by regular grammars).

We can make use of these theorems in the following way. Let us represent, as before, a (potentially infinite) set of leaves produced on a certain plant by a set of linear arrays and consider it as an adult language of an L-system. Then, if we can prove that the set of adult leaf-margins is not a regular set, it could not be the adult language of a right-linear OL-system. Similarly, if we can prove that a set of leaf-margins is not a context-free set, then it could not be the adult language of any OL-system.

These considerations result in the surprising finding that sets of symmetric entire leaves (without lobes—see class L_1 below) or sets of leaves with two identical lobes (see class L_2) can be generated as adult languages of OL-systems, while sets of leaves with three or more identical lobes (see classes L_3 and L_4) cannot be. A set of entire or two-lobed leaves can be described by expressions like

$$L_1 = \{kj^n mj^n k | n \geqslant 1\},$$

or

$$L_2 = \{kj^n kj^n k | n \geqslant 1\},$$

while a 3-lobed leaf corresponds either to

$$L_3 = \{kj^n kj^n kj^n k | n \geqslant 1\},$$

or to

$$L_4 = \{kj^n mj^n kj^n kj^n kj^n mj^n k | n \geqslant 1\},$$

depending on whether the lobes have tips (designated by symbol m) or not. As before k's stand for notches and j's for segments or serrations of lobe margins. Examples of the four classes of leaf shapes designated by the languages L_1 to L_4 are shown in Fig. 4.

It is well known by formal language theory that L_1 and L_2 are non-regular but context-free languages, while L_3 and L_4 are non-context-free languages. Thus, by the theorem of Herman & Walker, L_1 and L_2 can be adult languages of OL-systems, but cannot be adult languages of right-linear OL-systems, while L_3 or L_4 cannot be the adult language of any OL-system.

To illustrate this statement, we can easily produce an OL-system of which the adult language is L_1. The productions of this system are as follows:

$a \to kbk$, $b \to jbj$, $b \to jmj$, $j \to j$, $k \to k$, $m \to m$, and the starting symbol is $\underline{a}$. Since this is a non-deterministic OL-system, we obtain the derivations shown in Fig. 7 (the adult arrays are underlined).

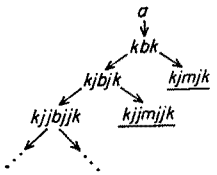

FIG. 7 Derivations of a non-deterministic OL-system of which the adult language is L_1.

In order to produce L_3 or L_4 one would have to have cellular interactions, or table OL-systems. How the latter kinds of systems can produce three-lobed adult leaves (for instance of the type L_4) was already demonstrated in the previous section.

As an illustration of the generation of L_3 by an L-system with interactions, let us take the following $\langle 1, 0\rangle$ L-system. Let the set of state symbols consist of letters a, b, c, h, j, k, let the starting symbol be a, and let the productions be as follows: $a \to kbk$, $b \to jck$, $c \to jc$, $c \to k$, $k \to k$ if the left neighbour of k is j or k or the environment, but $k \to jkj$ if the neighbour is c or h; furthermore, $j \to j$ if the left neighbour is j or k, but $j \to h$ if the left neighbour is c or h; finally $h \to h$ if the left neighbour is c or h, but $h \to j$ if the left neighbour is j or k.

Since there are two productions for c independently of its left neighbours, this is a non-deterministic $\langle 1, 0\rangle$ L-system, and we obtain the derivations shown in Fig. 8 (the adult arrays are again underlined) beginning with symbol a, and assuming a constant left environmental input.

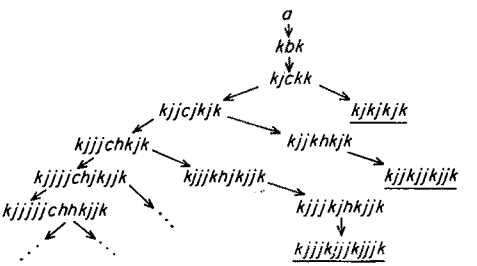

FIG. 8. Derivations of a non-deterministic $\langle 1, 0\rangle$ L-system of which the adult language is L_3.

Clearly the adult language of this system is the set $\{kj^nkj^nkj^nk | n \geqslant 1\}$ which is our set L_3. For every n, the generation of an adult array with n-times j

symbols in each segment requires $2n+1$ derivation steps, which seems to be minimal for sets with three equal segments.

Thus we are permitted to conclude that sets of adult leaves with three or more equal-sized lobes (or with lobes which bear a constant size proportion to each other) cannot be produced by development without cellular interactions. Sets of adult leaves with one or two symmetric lobes can be generated without cell interactions, but not by systems in which cell divisions generate a further dividing cell only to one side (so called right-linear OL-systems). The generation of sets of adult leaves with three or more lobes requires either interactions, or some form of global control over the cellular arrays, such as represented by table OL-systems. We do not believe that this connection between the generation of adult stages of organisms and the types of cellular processes required has been pointed out before.

7. Growth Functions

A significant and fruitful question regarding cellular developmental algorithms has been that of their growth functions. For each deterministic L-system S the growth function $f_s(n)$ is defined to have an integer value which equals the number of symbols in the array generated by S at the nth derivation step. Thus, the deterministic OL-system with productions: $a \rightarrow a$, $b \rightarrow ab$, $c \rightarrow bc$, and with an axiom consisting of 2 a's, 5 b's and 2 c's, is associated with a square growth function, i.e., for this system we have $f(n) = (n+3)^2$. It is, in fact, quite easy to find deterministic OL-systems for many simple polynomial type integer functions.

Recently, the theory of growth functions of L-systems has received considerable attention (cf. Paz & Salomaa, 1973; Herman & Vitányi, 1974). Some of the mathematical results have biological implications.

For instance, no cellular developmental algorithm (L-system) can have a growth function which rises faster than exponential with number of derivation steps (time units).

Every deterministic developmental system without interactions (DOL-system) has a growth function of the form (for some k)

$$f(n) = \sum_{i=1}^{k} p_i(n)c_i^n,$$

where each $p_i(n)$ is a polynomial function of n, and in each exponential term c_i^n the c_i is a constant. In other words, all DOL-systems have growth functions which are sums of polynomial and exponential terms. From this theorem it follows immediately that no filamentous organism without interactions can grow according to a function which increases slower than any polynomial

function (with finitely many terms) but is not bounded by a constant. Functions which increase slower than any polynomial but without bound are, for instance, the logarithmic function and functions which contain terms with fractional powers of the number of time steps. Thus no interactionless organism can grow logarithmically or according to the square-root of time.

Developmental systems with interactions (IL-systems) can be found with a much wider range of growth functions, among them the ones mentioned above. For example, to obtain a familiar *S*-shaped growth curve which approaches an upper limit, one can construct various simple systems with interactions. However, these growth curves can also be obtained with interactionless systems, since they do not grow without a bound.

8. Conclusions

To sum up, we have indicated the underlying assumptions and the biological justifications for adopting certain kinds of discrete constructs to describe and gain understanding of development. These constructs, growing cellular arrays of finite automata, have proved to be rather stimulating and useful when one wishes to consider the growth and differentiation of filamentous, multicellular or multicompartmental organisms and may eventually be extendable to organisms with two- or three-dimensional descriptions. Biological insights may be gained from these models, on one hand, concerning broad classes of developmental mechanisms (development with or without cellular interactions; symmetric versus asymmetric interactions—polarity; the role of changing environment and of cell death in development; simple or interlocking cycles of states—compound structures, etc.), and on the other hand by providing us with a framework within which we can construct specific models to simulate the development of certain organisms in detail. The developmental systems we propose here are able to provide only sketchy and tentative models, and they can certainly be improved upon. Some important developmental processes such as cell movement, cannot yet be expressed within these systems. Nevertheless, unexpectedly gratifying results have been obtained so far, and we are hopeful that this trend may continue in the future.

REFERENCES

BAKER, R. & HERMAN, G. T. (1972). *J. Bio-Med. Comput.* **3**, 201, 251.
CULIK, K., II & LINDENMAYER, A. (1974). Techn. Report CS-74-18, Faculty of Mathematics, Univ. of Waterloo, Ontario, Canada.
EHRENFEUCHT, A. & ROZENBERG, G. (1973). *Infor. Process. Lett.* **2**, 70.
ERICKSON, R. O. & SAX, K. B. (1956). *Proc. Am. philos. Soc.* **100**, 487, 499.

FRIJTERS, D. & LINDENMAYER, A. (1974). In *L Systems* (G. Rozenberg & A. Salomaa, eds) Lecture Notes in Comp. Sci. No. 15, pp 24–52. Heidelberg: Springer Verlag.
HERMAN, G. T. (1970). *J. theor. Biol.* **29,** 329.
HERMAN, G. T. (1971). *Int. J. Systems Sci.* **2,** 271.
HERMAN, G. T. (1972). *Int. J. Systems Sci.* **3,** 149.
HERMAN, G. T., LINDENMAYER, A. & ROZENBERG, G. (1975). *Math. Systems Theory* **8,** 316.
HERMAN, G. T. & LIU, W. H. (1973). *Simulation* **21,** 33.
HERMAN, G. T. & ROZENBERG, G. (1975). *Developmental Systems and Languages* Amsterdam: North-Holland.
HERMAN, G. T. & VITÁNYI, P. M. B. (1974). Techn. Report 82, Dept of Comp. Sci., State Univ. of N.Y. at Buffalo; to appear in *Am. math. Mon.*
HERMAN, G. T. & WALKER, A. (in press). *Int. J. Comput. Math.*
LINDENMAYER, A. (1968). *J. theor. Biol.* **18,** 280, 300.
LINDENMAYER, A. (1971). *J. theor. Biol.* **30,** 455.
LINDENMAYER, A. (1974). In *L Systems* (G. Rozenberg & A. Salomaa, eds) Lecture Notes in Comp. Sci. No. 15, pp. 53–68. Heidelberg: Springer Verlag.
LINDENMAYER, A. (1975). In *Developmental Systems and Languages* (G. T. Herman & G. Rozenberg, eds). Amsterdam: North-Holland.
LÜCK, H. B. (1974). In *Proc. Conf. on Biologically Motivated Automata Theory, held at McLean, Va., June* 1974, pp. 77–80. Long Beach, Calif.: IEEE Computer Soc.
PAZ, A. & SALOMAA, A. (1973). *Info. Control* **23,** 313.
ROZENBERG, G. (1973). *Info. Control* **23,** 357.
ROZENBERG, G. (1974). *Discrete Math.* **7,** 323.
ROZENBERG, G. (in press). *J. Comput. System Sci.*
ROZENBERG, G. & LINDENMAYER, A. (1973). *Acta informatica* **2,** 214.
SALOMAA, A. (1973). *Formal Languages*, part 2, section 13. New York: Academic Press.
SURAPIPITH, V. & LINDENMAYER, A. (1969). *J. gen. Microbiol.* **57,** 227.