T H E   U N I V E R S I T Y   O F   M I C H I G A N

COLLEGE OF LITERATURE, SCIENCE, AND THE ARTS
Computer and Communication Sciences Department

Technical Report

FORMALISMS FOR LIVING SYSTEMS
(Part I)

Richard A. Laing

ABSTRACT

Various formalisms for expressing the structure and behavior of biological systems are proposed and discussed. The formalisms include tiling or "domino" systems, the von Neumann cellular automaton system, and several systems based on the monogenic normal systems Tag and Lag. There is some discussion of desirable characteristics of models for biological theory.

CONTENTS

1. INTRODUCTION

This paper is "non-standard" in both form and content, so some explanation seems in order. First of all it is an <u>essay</u> (in the sense of an attempt, a tentative appraisal) intended to be exploratory in nature. In addition, what is presented here is only the first part of what is expected to be a series of connected expository and ruminative papers on the relationship between automaton theory and biological theory.*

In this present paper we show that certain of the results we should wish to obtain, may be, in a precise mathematical sense, <u>not</u> obtainable. This should not be allowed to become the occasion for deep pessimism. The negative results herein indicated have the value that by them we may begin to see some of the limitations of our models, and perhaps gain some hints as to directions in which the models may have to be modified in order to become more relevant to problems in "real" living systems.

A discussion of the limitations of the present study, and some thoughts on future directions to be taken will be found in Section 6.

Some brief remarks on the general relevance and the utility of attempts to relate formal systems and living systems are perhaps in order at this point.

I believe that there is mutual benefit to be derived from developing more closely the connections between formal systems and automaton theory on the one hand, and living systems on the other. Formal systems and automaton theory can serve as the vehicle by which the large body of results, not only in these two fields, but in algebra, graph theory, combinatorics, etc. can be brought more directly to bear onproblems of living systems. Is

there anything that biology can offer automaton and formal system theory? The answer certainly is yes. The immense complexity of living systems, the large numbers of concepts and of important open problems can serve as an extremely valuable source of "directions" in automaton theory and formal system theory. Indeed, it is characteristic of automaton theory in particular that it acquires new directions by the contemplation of problems of the real world (for example, computer programming or technology) where its applications are thought to lie. The time is ripe for the concepts and problems of living systems to become a major source of new directions in automaton and formal system theory.

[It is not my intention to denigrate the work that already has been done in connecting living systems and formal or mathematical disciplines (e.g. the work done in the "cybernetical-control-bionics areas") but to attempt to enlarge the area of contact and accelerate the exchange of ideas.]

One more point should perhaps be made at this time. Much of what is said in this paper may seem more than a trifle strained or unreal to a working biologist who feels his field is defined in terms of "real" carbon molecule chemistry. I am going to try not to be embarrassed by this.

I can only say: I am beginning by contemplating broad biological-like processes in possible worlds. By suitably specializing and restricting the formal or automaton systems, one should be able (perhaps by a lengthy process of approximation) to come at last to a "good fit" with the real living systems a biologist would recognize and be directly interested in. Activity in this direction, I think, will readily be acknowledged as valuable.

On the other hand, I think there is also a definite value in specializing

and restricting very general systems for biology to yield living "artifacts"--
artifical plants and animals which are <u>not</u> like those we can be certain to
find in our earthly (or perhaps any other) garden.

This design of artificial systems can be justified (if justified it
needs to be) by suggesting that it is a reputable mathematical activity in
which the customary criteria of taste, elegance, economy, relatedness, etc.
ought to serve as guide and arbiter. Also it is good practice in sharpening
one's old tools and devising new tools for the problems of the "real"
biological systems; also somewhere in the universe these models of artificial
plants and animals, may have <u>real</u> counterparts in plants and animals.

2. MOLECULAR AND CELL-LIKE SYSTEMS. SYSTEMS OF CELLS AND MOLECULES.

Let us now describe, in broad terms, the sort of "living systems" to be considered here.

This description is rather free-wheeling, and intended to be more suggestive than precise. The later sections of this paper illustrate what is meant, and by this means "inductively" provide a more exact sense of much of what is presented broadly in this section. The systems to be considered will, first, have elements which can be viewed as cell-like or molecule-like. These elements may have connective properties, so that certain of them may be able to cling together or adhere; they may have relational pro- perties in space, so that there are preferred orientations to other elements or to the "space" of the systems of which they are a part; they may have information carrying and transmitting properties, so that they "communi- cate" to other elements or aggregates of elements. The properties we have just mentioned permit the elements to assume structural relationships akin to the strengths, distances, and angles of chemical bonding. An aggregate of elements in a structure, may itself have adhering, relational, infor- mation bearing and transmitting properties just as the elements themselves. The principal difference between elements and structures of elements is that the structures may be decomposed, the elements not. The structures however, may be very stable and may only rarely or under particular condi- tions undergo change; on the other hand, the structures may have numerous regions or sites which have weak or unsatisfied or special bonding conditions so that they are susceptible to change at that point, or to the attachment of other structures at that point, or to the assembly or catalysis of elements or structures at that point. When either element or structure is meant, the neutral term constituent will be employed.

Some constituents will have a purely structural "scaffolding" role,

while some constituents may have roles both as determiners of particular

structural form and as direct agents in attachment, assembly, synthesis,

information transmission, etc. of other constituents.

These latter "active" agent constituents may have their activity highly

dependent on the structural properties they possess and on the structural

properties of the "non-active" constituents, since the form of a constituent

(the order, proximity, arrangement, strength, etc. of the related constituents).

may determine the relationships of active constituents∧ Thus constituents

having a "fundamental" (primary) chain-like form may have their activity

greatly affected by any folding or "switch-backs" which bring active con-

stituents into conjunction: thus active constituents in chain-like mole-

cules may be linearly distant but spatially and functionally adjacent.

[For example, the hemoglobin and myoglobin molecules, though "funda-

mentally" linear, have a three-dimensional structure which creates active

enzyme production sites formed of linearly distant elements.]

In the systems we wish to model, we will customarily ignore the problem

of specific energy requirements for chemical action (that is, energy

requirements will not enter separately and explicitly into our cogitations)

and may at times also assume that chemical materials of various sorts are

freely available; assembly, binding, breaking and transport of constituents

will be assumed to take place either "spontaneously", or as a consequence

of specific rules of action. Thus, if we should wish to model the behavior

of a DNA molecule, we might assume that thermal activity "freely available"

was sufficient to break the hydrogen bonds connecting the separate DNA

strands, while we might either assume that "free-floating" nucleotides and

bases would be available to the exposed active sites, or on the other hand,

that we must model a larger system which contains the production sites pro-

viding the required nucleotides and base materials. In either case, we

should probably be most interested in embodying the information bearing and

and transmitting apparatus, that certain complementary pairing rules have to be obeyed in order to form new DNA strands, and that these information carrying strands can separate from the parent strands to take part in further replication or in protein synthesis. (In effect we will try to connect (as far as it is possible) problems of material and energy to problems of rules and information.)

Still speaking generally then, we will be considering structures of elements, these structures being capable of producing elements or structures of elements, and of attaching themselves or being attached to other structures, the "fluid" medium in which these constituents exist assumed (if necessary) to contain the material and the energy to permit the assembly, binding, breaking and transport of constituents.

In this paper, we focus attention principally on various systems of various rules for the production and attachment of constituent structures, and on the ensuing consequences of these rules. In the next section we examine some systems in which the constituents are assumed to be freely supplied and the rules specify which attachments may be made.

3. CONNECTION SYSTEMS

In this section we consider systems in which we are given as "raw material" sets of constituents which we may view either as molecule or as cell-like structures. These constituents have connective properties which allow them to be formed into larger structures. The connectivity constraints are the rules of the systems, and we wish, for various classes of rules, to decide, for example, what structures can or can not be formed. We begin with a linear, double-stranded system of complementary constituents (resembling a non-helical DNA), and later we consider a "membrane" constructing system.

3.1. A DNA-Like Double Stranded Linear System.
Let us suppose that we have available to us macro-molecular-like constituents of the sort, two examples of which, are pictured below:
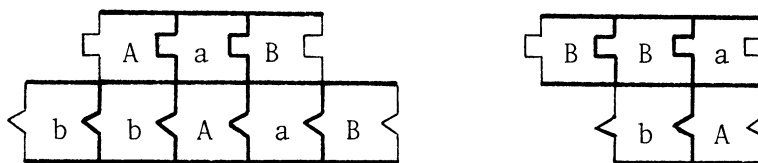
Fig. 3.1.

That is, we are given fragments of a double stranded linear system, in which "tongue-and-groove" connections along the upper and lower strands are meant to suggest bonds of considerable strength, while the connections between the two strands are pictured as shared surfaces, suggesting by this rather weak bonds, so that breaking and "sliding" and reforming along this line of bonds is assumed to be readily possible. The kinds of constituents of the upper and lower strands are the same; for each of these kinds of constituents in one strand, there is a single kind of complementary constituent

in the opposite strand. (We may express this by employing upper- and lower-case letters; the upper-case letters have as their complements the corresponding lower-case letters, so that there is one match and two mismatches in the first example above, and two matches in the second example above.)

The given double strand fragments may not (as is seen in the first examples) have all their elements properly paired; such mismatched double strands are less stable and more susceptible to bond breakage and sliding. The more elements properly complemented, the more stable our double-stranded structure will be. If, throughout the length of a double strand, all the elements are properly complemented, we say that the strand is completed.

The following question would perhaps have some interest to biologists: Given a set of such double strand fragments (and unlimited copies of each kind of fragment) is it possible to tell whether completed strands can ever be formed? The answer to this is: No, it is not possible to tell in all cases.

[This result is based on the recursive unsolvability of the Post Correspondence Problem. In the Post Correspondence Problem, a finite set of word-to-word dictionary-like entries is given and the question asked whether one can always tell whether a selection of dictionary entries can be made which will spell out the same sentence (letter for letter) in both the source and the target languages. Post himself showed that in general it is not possible to tell (when the alphabet on which the words are formed has more than one letter). Our upper and lower strand fragments form, e.g. the source and the target language entries; our completed strands are Post's correspondences.]

3.2. A Membrane Construction System. Suppose we are supplied with unlimited copies of a finite number of different constructional units from which we can form cell surfaces, grids, sheets, membranes, etc. Other

instances are sheets formed of cells, or alga colonies on a fluid's surface. The units have some definite orientational properties: they have an "inside" direction and an outside (so that for example, preferred directions of membrane porosity might be implied) or an "up" and a "down" orientation. The units also must be in proper alignment for connection to take place, as well as possessing appropriately complemented bonding forces.

Thus, in the diagram of Fig. 3.2, a specific orientation is implied or no "fit" will be possible (regularities will be severely disrupted); in addition adjacent units must possess complementary "adhesive" values: the adhesion properties right, left, above, and below (or, as viewed from above as in a liquid surface, east, west, north, south) any given unit must match for the given unit to fit.

The question, "can we tell, given any set of such construction units, whether a surface of indefinite extent can be constructed?" is not decidable.

[This result follows from the unsolvability of the tile or domino problems: It is recursively unsolvable to decide whether given any set of unit sized square tile types with edge coloring (allowing translation of tiles but not rotation or reflection) observing edge-color matchings, the whole plane can be paved. Certain domino problems with constraints have been shown unsolvable (by Wang and others) while the unrestricted problem has been shown unsolvable by Berger.]
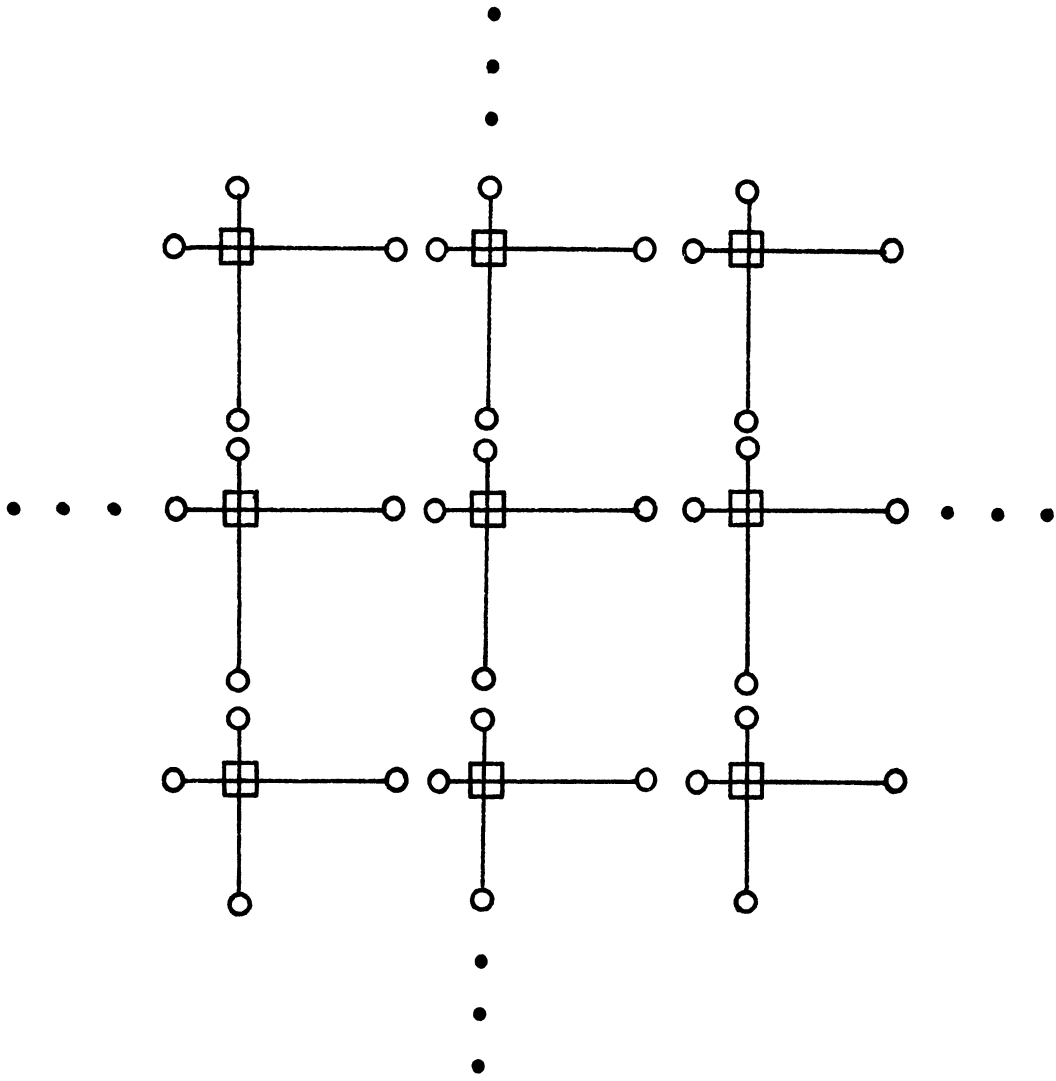
Fig. 3.2.

4. SYSTEMS OF PRODUCTION AS WELL AS ATTACHMENT

In the last section we considered some systems in which constructional units were incorporated into linear or into surface structures. We there assumed that we had available to us all the units which were to be fitted into the structures, and we did not concern ourselves with any problem of the source of these units. In this section we consider some systems which in a sense take part in the production of their constructional units at one site while at another site they incorporate the units into themselves. Thus there is a sort of "feed-forward", the production consequences of the present having possible production consequences in the future. The potential for certain mechanisms by which activity can be controlled (e.g. shut down or changed over to new production) is obvious.

4.1. A "Molecule-Like" System that Produces Molecular Fragments. Suppose we have given a "chromosomal strand", which has along one end an active production site. The site, in combination with the fluid medium in which the strand is immersed, will produce a fragment of new "chromosomal" strand, which, moving freely through the fluid, in which the process takes place, attaches itself to that end of the original strand opposite to the action site. As a consequence of production, an active site is deactivated and also as a consequence of production, a new site, located some fixed distance along the original strand, is activated and begins to produce the next linear fragment.

Thus, linear fragments are produced at active sites along the original strand, and are attached to the opposite end of the original strand in the same order in which they were produced. If the process continues long enough, the produced fragments (or portions of them) will themselves become part of the production determining process.

The process of production and attachment will continue as long as there remains a length of activated strand, and an activated site produces something. (If an activated site fails to produce anything whatsoever, then it is not deactivated, and does not activate the next site along the strand; at this sort of blockage activity ceases. This case should be kept distinct from the case where the deactivation and new activation takes place but no new fragment is produced for attachment at the attachment site.) In Fig. 4.1, one way in which this process might be carried out is shown. Note that secondary structure of the linear system could bring the active production and the attachment sites into close juxtaposition, facilitating formation of the completed strand.

Can we tell whether in such a system, activity will go on indefinitely? As in our earlier cases, the answer is, "No".

[The underlying system here is a monogenic normal system of the "Tag" sort; in Tag, the first letter of a given string determines a word to be attached to the opposite end; a fixed number of initial letters are then deleted (deactivated), so that there is a new (activated) "first" letter, which determines the next word to be appended, etc. "Tag" (the halting of the process) takes place when deletion catches up with production (or when an initial letter is not the antecedent of any production rule). That the halting problem for Tag is unsolvable was shown by Minsky. The "word problem",--whether you can tell whether some particular string will ever be produced--is also unsolvable; it also follows that the problem of deciding whether two TAG systems produce the same string is unsolvable.]

4.2. An Artificial Plant. To illustrate further the uses to which monogenic normal systems may be put we present an "artificial plant" interpretation of the system.

Our plant begins with a "seed" which produces an initial "jointed vine", each "joint" being composed of a single cell-like unit. The initial growth from the seed having been completed, further poant growth will be a function of the characteristics of root-end sequences of cell-units. These root-end

Connection-
Growth
Site

Newly Produced
(Freed) Constituent
Will Attach To
Growth Site

Deactivated
Region

Active
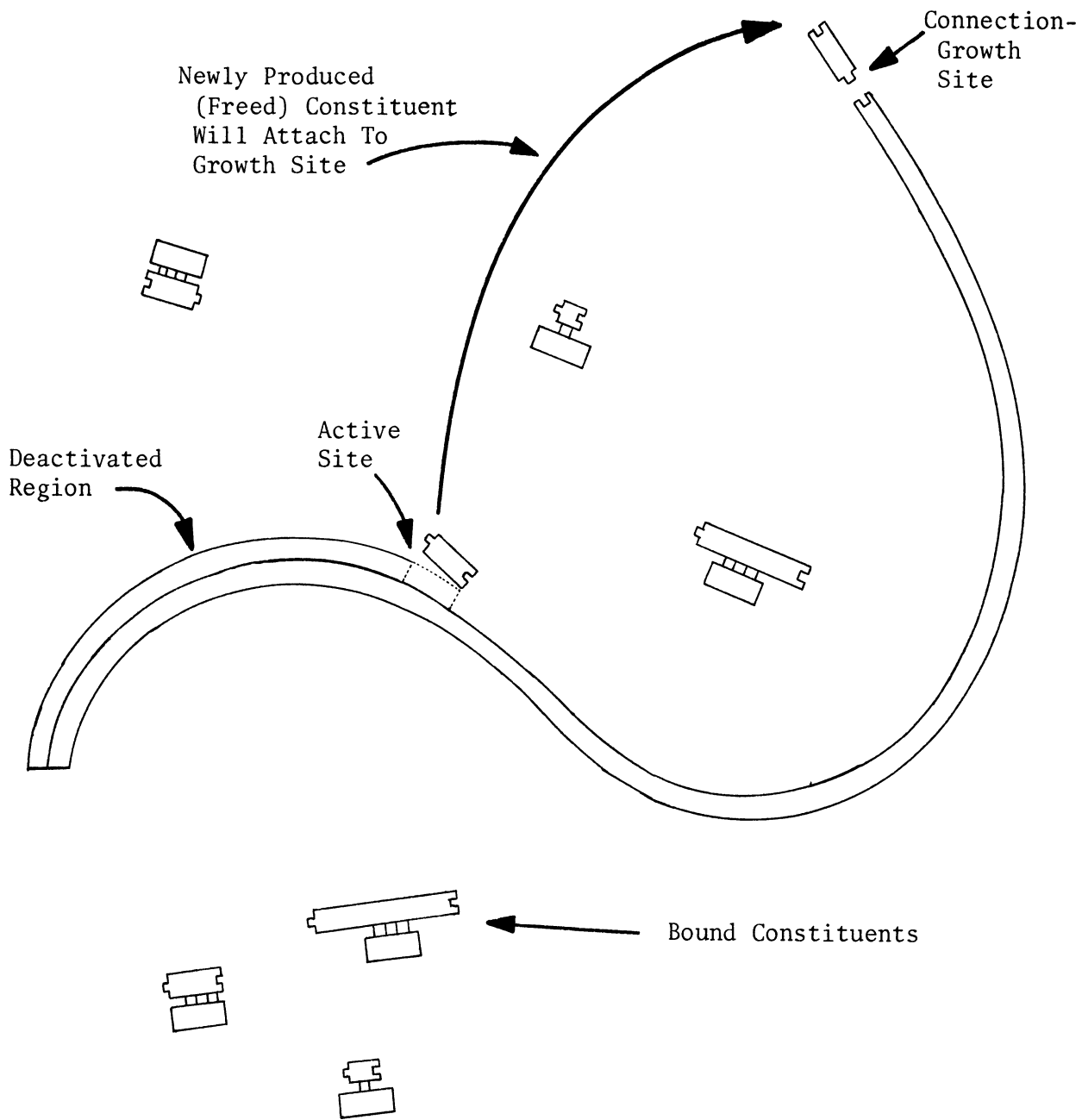Site

Bound Constituents

Fig. 4.1.

cells have a generative function and for this reason we call them the

<u>mature</u> <u>cells</u> of the plant. Any mature cell will as time goes on, eventually

cease to have any generative function and will harden into purely struc-

tural or connective plant tissue. We will call these latter cells

<u>senescent</u> or <u>aged</u> <u>cells</u>. New cells will be produced at the end of the

plant opposite from the root. New cells produced at the growing tip of

the plant will be called <u>immature</u> <u>cells</u>. Any cell which is neither

mature nor aged will be immature. Each of the three classes of cell types

will be contiguous along the plant tendril; that is, taking them in order

from the root end, these would be <u>aged</u>, <u>mature</u>, <u>immature</u>.

Let us force the plant or flower analogy a bit more and say that the

cell segments of our artificial plant can assume any of a finite number

of colors. We might say that cells, as they pass from the mature stage

to the aged stage, go from some bright flowering color to gray; the immature

cells need not exhibit their colors openly, until they mature, but what

these colors will be will have already been determined through their mature

cell parents. This "paleness" of the immature cells can be viewed as

having a functional counterpart in that they will be "transparent" to

growth propagations from the mature segment, through them, to the growing

tip of the plant. That is, the genetic effect of the mature segment of

the plant will pass through immature cells, to the growing tip, without

for example, changing the color heredity of the immature cells. The

sequence of color qualities exhibited by the mature portion of the plant

we may assume to be a direct expression of the underlying generative proper-

ties of that segment; by this color sequence we can determine the eventual

colors of the new cells, growing at the tip, as a consequence of the generative

action of the mature segment. (As described here, there is a one-to-one

correspondence between underlying genotype and the expressed color pheno-

type; we could of course add to the complications by employing "false"

colors, etc.)

As is customary with plants and other living things, growth will cease

whenever aging consistently takes place faster than new cells are formed

What can we say of our plant? Can we tell whether a certain pretty

color pattern will ever appear? Whether the plant will grow "excessively

large"? Will it cease to grow soon?

Let us describe the growth of the plant more carefully. At each

growth phase of the plant, the heredity (expressed outwardly by color)

of the first k (where k is some fixed finite number) of non-aged cells

will determine the precise numbers (perhaps none) and color-kind of cells

to be formed at the growing tip. (The first k non-aged cells are the

mature cells of the plant.) The growth impulse is transmitted through

the immature cells to the growing tip. At the conclusion of the growth

phase the oldest mature cell ages, and the oldest immature cell matures.

Each mature cell may thus take part in as many as k growth phases. If

by this process k mature cells should not be available to initiate a

new growth phase, all further growth ceases.

The following can be shown: If the plants have a mature portion at

least two cells long then there is no effective process which will tell

for all seeds whether plants of this sort will ever cease growing or

produce some particular arbitrary "pretty" color pattern.

These results are a consequence of the properties of the formal

system, which (just barely) underlies our plants.

[In this instance the system is the so-called "LAG" system which has

universal computational properties and thus unsolvable halting and deriva-

bility problems. (The fundamental undecidability result is by Wang (1963)

and was proved by showing that the action of any Shepherdson and Sturgis

Machine (1960, 1963) can be expressed in a LAG system.)]

5. CELLULAR SYSTEMS FOR AUTOMATON GROWTH
AND SELF-DUPLICATION

None of the formalisms thus far introduced had its genesis in an
attempt to model living systems. They were all originally devised to
explicate some aspect of computational or algorithmic processes. It may
be felt that the genesis of these systems accounts for the biologically
"curious" unsolvability results, and that formalisms devised with living
processes as the interpretation in mind would be less likely to yield
such oddities. In this section, we briefly indicate how two such systems
purposely designed to model interesting biological characteristics also
have unsolvable properties associated with them.

5.1 The von Neumann Cellular System for Growth and Self-Duplication.
John von Neumann designed a cellular automaton system for which he proved
that machine self-reproduction was possible. In his system, both parent
and progeny are constructed from a single set of cell types, each type
being representable by the state of a single kind of finite automaton.
The machines, parent and child, can be of indefinite complexity; indeed
machines capable of universal computation are constructible. His machines
"live" in a two-dimensional checkerboard-like array. Each square of the
array houses an identical 29-state finite automaton. The precise state
(cell type) of one of these automata is a function of the immediately
previous state of each of its four cardinal direction neighbor automata
as well as its own previous state. The precise state a cell automaton is
in determines whether it sends signaling or construction-destruction pulses,
what sort of logical switching role it plays, the direction of its activity,
etc. (Details of the von Neumann system may be found in papers by Thatcher
and in Burks' completion of von Neumann's design; knowledge of the precise

details is not necessary for what follows here.)

Von Neumann showed that some of the machines of his system are capable of self-duplication. That is, that beginning with a particular single machine in this two-dimensional space, after some period of time there will be a second machine, identical to the first, and this second machine will itself proceed to make a copy of itself, etc.

Some machines of course will not be self-duplicating machines. There will also be machines for which it will not be clearly apparent whether or not they are self-duplicating machines; that is, for example, they may seem actually to start on the job of constructing a copy of themselves but then abandon it or seem to abandon it.

Can one tell in general whether any given von Neumann cellular automaton is a self-duplicator? Once again the answer is "no". Universal Turing computation is possible for a von Neumann cellular automaton. Of all the von Neumann cellular automata, there will be an infinite number of those possessing Turing machine computing properties. Indeed of all the von Neumann cellular automata there will be many (in fact infinitely many, though this is not important) each acting like any given Turing machine. Of these there will be for each kind of Turing machine, machines which will duplicate just in case their associated Turing machine halts. That is, for example, there will be a class of von Neumann cellular machines possessing "sufficient" self-duplication "equipment", but these machines will not actually duplicate unless their associated Turing machine halts and transmits an "initiate duplication" command signal. Since Turing machine halting is known to be undecidable, self-duplication in this general sense, is also undecidable. (Notice also, that although von Neumann cellular machines are capable of performing all sorts of

interesting and complicated tasks, we can employ the same reasoning to show that there can be no procedure by which we can tell in general whether the cellular machine <u>will</u> perform the task.)

Thus even for a system purposely designed to possess interesting biological-like properties there are characteristics which may seem to the biologist "pathological".

<u>5.2. A Cellular String System</u>. The unsolvability results that arise in the von Neumann cellular system are a direct consequence of von Neumann's intent that his system embrace a "universal" degree of functional-computational complexity; he did this (among other reasons) so that no accusation of triviality might be leveled at his self-duplication results. That is, the very thing which produced the unsolvability consequences in the earlier computation systems, has been deliberately embraced in a system designed to model a biological-like activity. Is there a formalism, designed for biological-like activity, in which the designer has not deliberately incorporated universal computation?

Lindenmayer has devised a cellular automaton system consisting of (a) an array of cells, each of which houses a finite automaton, (b) an environmental input consisting of a periodic series of impulses of various kinds which can be injected into any of the cells of the array, (c) a set of rules by which upon receipt of input impulses from environment or from neighboring cells, a cell automaton may change states, produce new cells (which may occur in such a position as to extend a line of cells, separate two cells, or initiate a branching off from an already formed line of cells, etc.).

Lindenmayer shows how his system can produce linear arrays of cells modelling filamentous organisms with constant apical patterns, banded patterns, etc.; he also shows how his system can produce branching arrays

of cells resembling the structure of a type of red algae.

[Lindenmayer's system closely resembles the "growing automata nets" of M. J. Apter. The most striking immediate difference between these two formalisms for biological systems undoubtedly arises out of a difference in intended application. Lindenmayer models "filamentous organisms" and consequently allows new cells to arise not only at the ends of his filaments but between already existing cells of the filament as well. On the other hand, Apter, since he assumes less pre-existing structure than Lindenmayer, turns every new cell into an incipient branching process; that is, a cell can not arise, in line, between two cells already in line, but is grown in adjacency to its parent only. Most of this "natural"-ness in Apter's system is, however, lost by his use of a tape communicating between automata. This tape seems to be an unnecessary and awkward feature, undoubtedly left over from the Turing machine source of some of his automaton ideas. Since it is not entirely clear what Apter would permit in his growing automata nets, it is not clear whether his system has universal computation properties. The fixed directions of information flow (along the tape) certainly constrains what can be done, but as we have seen in Section 4.3 Universal Computation in a cell system with information flow in one direction is not absolutely crippling.]

For example, in producing an apical-like cell array, Lindenmayer employs cells limited to the two states 0, 1, and which have input impulses from the "root-end" environment, which consist entirely of 1's. The cell automata outputs are identical to their internal states, that is, a cell automaton in internal state 1 will have an output 1, and a cell automaton with internal state 0 will have an output 0. The cell automaton outputs are, in this particular case, always directed away from the root-end and

toward the apical end of the cell array. The specific cell rules are the

following: if the cell is in state 1 and receives a 1, it goes to state 0;

if in 1 and receives a 0, it stays in state 1 and also duplicates itself

(producingtwo cells in state 1): if in state 0 and receives a 1, it goes

to 1; if in state 0 and receives a 0, it remains in 0.

If we apply these rules to an initial linear array of five cells in

states 01100, a stable pattern is generated, moving from the left to the

right, although the cells making up the pattern continually change state

and are displaced.

At least in its obvious intent, Lindenmayer's system is more bio-

logical-like than von Neumann's; many processes much like those found

in living systems can be exhibited; perhaps more important, the cell

arrays are not locked in a rigid discrete geometry such as von Neumann's

"checkerboard": cells can arise between cells, and three dimensional

shapes can be formed. (Lindenmayer is aware of, but does not address

himself to the problems of location, crowding, etc. which necessarily

arise when many new cells arise from a single basal cell.)

Even such "natural" systems as these however are subject to the

same troublesome solvability constraints we have noted so frequently above.

Given any Turing machine, a linear array of Lindenmayer-system

cell-automata can be found such that the sequences of internal states

will (in a suitable fashion) correspond to the sequences of tape configura-

tions of the Turing machine. It is well known that any Turing machine

process can be realized by some Turing machine that begins with a blank

tape, and that employs for printing on the tape only two symbols. (These

results are not at all necessary to the proof, but they make simpler the

language of this proof-sketch.) Our argument then goes like this: we

begin with a single cell automaton which is basically the same as the
finite control-head automaton of the Turing machine to be imitated. Our
cell-automaton will have some additional properties (all allowable in
Lindenmayer's systems). Some additional (beyond the Turing state table)
part of its internal state configuration will record that it has the tape
value 0 or 1; some part of its internal state configuration will record
that it is or is not the active control cell in the array; also the cell
will be able to produce copies of itself, on the right or left (although
again previous results would allow us to confine ourselves to one side only);
finally it will be able to send and receive state-changing pulses to and
from cell neighbors to the immediate right and left.

The initial cell automaton will have a tape value (i.e., 0 to begin
with) and will be in the active mode. If its internal Turing state table
calls for e.g. the "printing of 1 and a move right", then this initial
cell will go into the internal configuration standing for the recording
of the tape state 1, will construct a copy of itself to the right, trans-
ferring control to this new cell which will also be in the internal config-
uration it will be placed in the internal state the old active head would
standing for "looking at a tape 0"; for the rest of its internal configuration
be in after the print and move right had taken place. If the next Turing
machine action should be e.g. a "print 1 and move left", then our newly
created cell would go into the tape value 1 state, and then transfer both
active control and the new Turing-table internal state information back
to the original cell.

It should now be clear how to proceed.

Since we can not tell, given any Turing machine whether it will
halt, we can not tell, given any Lindenmayer system, whether it will con-
tinue to grow. (Notice that we have not needed to employ either stimulus

pulses from the outside, or array branching rules. We have, however,

employed two very powerful properties: the system permits finite cell

automata of arbitrary complexity, and allows duplication of automata in

a single operation.)

5.3. A Related String System. Another system (let us call it L_2)

closely related to that of Lindenmayer, may be of interest. In L_2, we will

permit non-deterministic cell automaton behavior, in the sense that the

precise cell behavior (internal state change and any consequent growth

or "signalling") is to be arbitrarily selected from a finite repetoire of

possibilities.

[This non-determinism is not essential to what is being discussed

here; however, the assumption of non-determinism greatly simplifies the

description of certain phenomena. Non-determinism could be eliminated by

introducing a "selector" input signal from the environment which determines

which element of the repetoire of possibilities is to be chosen, or, in

the case at hand, a deterministic cell system could be appended to a cell

system of our L_2, which would systematically take the non-deterministic

L_2 system through every possible action in its repertoire. Cf. Dana Scott

on non-deterministic vs. deterministic machines.]

Let us allow a cell-automaton, as a consequence of particular

internal states, to produce sequentially a finite string of new cells;

these offspring cells need not be completely identical to their parent

cell.

Then we can say that it will be impossible to tell in general

whether two organisms of the L_2 type (given by their growth rules, etc.)

will have identical cell sequences.

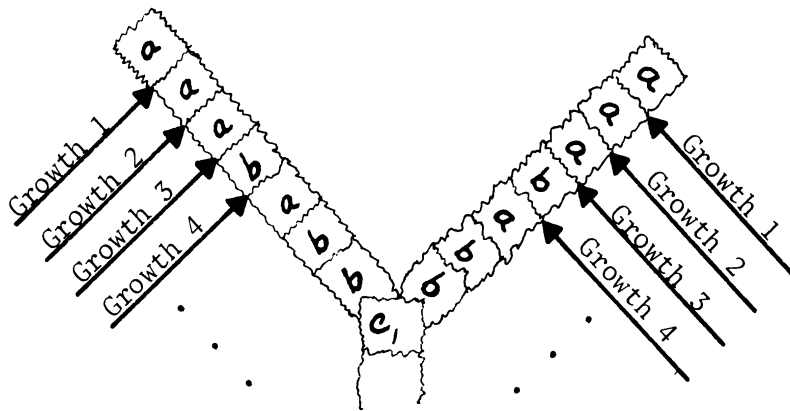(The proof, briefly sketched, is this. Take as the first organism,

O_1, of L_2 the following: the organism O_1 will begin with a single basal cell c_1 (c for "center") and immediately branch. This c-cell will have the property that whenever it produces a cell at the left branch, an identical cell type is produced at the right branch. Two produced cell types are sufficient for our argument; let us call them a-cells and b-cells. Any pair of branches of our O_1 system will thus, throughout its life, have the same a-cell, b-cell pattern.

Take as O_2, our second organism of L_2, the following: O_2 is also a branching organism beginning from a single basal cell c_2; c_2 will have a finite repertoire of deterministic sub-routines; each of these subroutines will cause c_2 to produce some finite string of a-cells and b-cells for the right branch, and some (usually different) finite string of a-cells and b-cells for the left branch. See Figure 5.1.

The pairs of branches that an O_1, type organism produces, and the pairs of branches that an O_2 type organism may be identical or different. If we could examine an O_1's and an O_2's rules and tell whether their pairs of branches will ever be identical we could tell if a correspondence can take place in any Post Correspondence System. We know we can not do this; the Post Correspondence Problem is recursively unsolvable. Therefore, we can not in general test organisms of L_2 for yielding identical patterns.

[This result would also follow from the universal computation properties already shown for Lindenmayer's system. The result presented in this correspondence system form should be immediately familiar to computational linguists and was couched in this fashion to suggest that many of the results of computational linguistics might readily be applicable to cell systems.

A valuable discussion of algorthmic unsolvability in a biological framework is given by Stahl.]

Organism O_1 (always produces same cells from c_1 for both branches)

Fig. 5.1.

Organism O_2 produces some finite set of cells along each branch; these need not be the same cell types nor of equal number.

Fig. 5.2.

6. DISCUSSION

We have presented some descriptive formalisms, and pointed out some of their characteristics. It is now time to look more closely at such systems, to try to determine their virtues and failings. First, however, it may be useful to consider what our goals are. Knowing some of the things we desire may provide us with some guidance in selecting desirable and undesirable characteristics in descriptive formalisms for biology.

6.1 Some Theoretical Goals. We may as well begin by setting our goals high, and asking for a great deal. We may of course have to settle for much less; but we will try to hold each line as long as we can.

It would be nice if there were a descriptive formalism adequate for all possible phenomena in all possible worlds. Some of the possible phenomena in some possible worlds, will be what we (here on Earth at least) call "living systems" or "organisms". Such a formalism would somehow embody all the laws of physics and chemistry, as well as the processes of mathematics and logic. (Physics may be said to subsume chemistry and logic mathematics, but these issues need not be pursued further. Whether we must somehow go beyond this, to account for all phenomena, everywhere, I do not know.) In effect we may wish to express phenomena in logically as well as causally possible universes.
[It has of course been frequently argued that "satisfactory explanation" of living systems must somehow go beyond the laws of physics and chemistry (to the "laws of life", the "laws of human behavior", the "laws of consciousness"). I personally believe that the laws of physics and chemistry are more than sufficient to account for living systems. For instance, I do not believe that subatomic phenomena has any bearing on the issues of biology.

Part of the continuing differences of opinion on this matter lies, I believe,

in what is meant by "satisfactory explanation". If by explanation is
meant some phrasing, some set of laws which serve readily to communicate
to other humans, I am in complete agreement that we must go beyond the "laws
of physics and chemistry", as these physical laws are usually expressed.
There should be combining, selecting, rephrasing, of the physical laws,
which apply to what we have selected of the world to call biology. I
thus think that it is a matter on the one hand of referring to the category
of events in the physical world we call biological events, on the one hand,
and of appropriately adapting one's style, one's rhetoric, one's pedagogy
to match predisposed categories of the human intellectual grasp. It is
not a matter of the inapplicability of the laws of the physical world (the
laws of physics and chemistry) but of selection and grouping of those
laws in accord with the subject matter as understood by a person when he
speaks of biology. Recently Polanyi argued against this saying that not
only are there levels of appropriate discourse, but that those levels signal
the real existence of separate realms whose laws are not the laws of
"lower" levels. (For Polanyi, there are "higher" and "lower" laws;
chemical laws are "low"; "spiritual" "laws" are "high".) I think the case
is rather analogous to the problem of founding all mathematics on logic
and/or set theory. Mathematics does seem to be "accountable" in terms of
logic-set theory. That the work-a-day mathematician handling differential
equations or problems of algebraic topology does not usually restrict
his thoughts and exposition to logic-cum-set theory does not mean that every-
thing mathematical he does is not accountable by those means but rather that
such meticulousness would be wasteful of time and energy, and it would
hurt the head to do it. Cf. Whitehead on symbolic notions and abstractions
being introduced to make things easier to think about. The notion of a "higher"
level being a simulation of the pertinent parts of the level below, is I
feel another valuable way to view the matter.]

As to particular research strategy, there is a whole range of possibilities. We could begin with "anything that is imaginable" and reduce this to anything which is consistent with our body of intuitive foundational logical and mathematical concepts. This in turn could be reduced to those things allowable in our physics and chemistry. In developing the relevance to biology there are many paths to choose from. One path is to consider all mathematical and logical processes; then to restrict one's attention to those which have physical analogues. Finally one can restrict attention to those physical analogues which we call biology. (Each of the above reductions would have to be discussed and justified.)

A second path (and the one we shall probably most often have in mind) is to move directly from mathematical processes in general to an important class which is biological-like (that is, to move directly to automaton theory). Next we consider which parts of automaton theory have physical counterparts. The arrived-at class should include everything we call biology (and perhaps many things which we do not call biology, but which are inextricably bound up in biological notions--computing artifacts, etc.)

The level at which the research here will be carried out is that of the logical and mathematical equivalents of the biological-like laws. That is, I wish to create and manipulate mathematical constructs rather than conduct experiments in physics or chemistry (or that part of physics and chemistry called biology). I do, however, wish to discipline my constructs in the direction of the physically realizable, so that the work may reasonably be said to have relevance to "real" phenomena.

I take it that biology is a sub-system of rules, expressed in the formalism, that the rules are discoverable, that the rules are "physical and chemical" in nature, and that they operate deterministically. I will

also argue that restricting attention to rules for discrete phenomena will be fruitful.

Of course, it would be nice if there were an algorithm which, for all phenomena, would distinguish the rules of living from non-living systems. Of course, if there _is_ (as is so often asserted) a continuity between living and non-living systems, then such an algorithm may not exist; we might then wish for an "algorithm" which will be of _some_ assistance in distinguishing systems. _Also_, if we are convinced of some special degree of complexity to be found in indisputed living systems, then we may very well require of the universal descriptive system that it be "especially rich" in descriptive power when a living system is under consideration. That is, we may wish to group the primitive concepts of the system to produce "combined" concepts easier to grasp and work with because they are closer to what we think of as biology.

Any particular organism or living system may very well be properly expressed in all its important features only if we recognize that its environment is part of the system. Can there be an algorithm which will tell for some indisputed "living" part of a large system, how much of the environment (in the broad sense) must be included along with the organism in order to account for all of its biological behavior of interest?

Perhaps this is enough background to allow us to be more specific about possible goals in biological theory. Let us therefore list some of these goals.

1) To establish a formalism adequate to describe (in terms of initial conditions and rules of behavior, say) all possible living systems in all possible worlds.

2) Given a set of initial conditions and rules, in the descriptive formalism, to tell whether the set defines an organism (living

system) in some world.

3) Given a set and a world, does the set define an organism in the world?

4) Given an organism (presented in some unspecified way) and (perhaps) its world, can initial conditions and rules be obtained?

(An organism could well be the consequence of several sets of initial conditions and rules--thus there may be an ambiguity of generational origin.)

5) Are there organisms which can be accounted for on the basis of several sets of initial conditions and rules? Only on the basis of a unique set?

6) Given an organism and a set, can one always tell whether the organism is a consequence of the set?

7) Given a possible world can all the organisms of it be enumerated.

(Since some of the algorithms implied above may not exist; weaker goals be appropriate--enumeration being weaker than decidability.)

8) Given an organism and two sets of conditions-rules, can the set which better accounts for the organism be established?

Many more questions and goals will undoubtedly occur to the reader (as they have to the writer) but perhaps this gives some of what might be asked of a biological theory.

[Some of the above problems would, under particular interpretations of their terms, collapse together into a single problem. Solutions to some would provide solutions to others. Some perhaps are nonsense.]

The working biologist would undoubtedly approach many of these questions by trying to conceive of tests, experiments, which would provide the answers. The automaton theorist will think in terms of finitely defined processes, algorithms, providing the answer. This relationship between

the two disciplines though presently very tenuous ought to be exploited

by both theoretical biologists and automaton theorists. The principal

general mechanism of co-operation will be this: the automaton theorist

will supply many ways in which life-processes could be carried out; this

is not to say that these will be in fact the way or ways in which the

processes are carried out. This latter matter is a problem of verification

in the laboratory. The research here proposed will supply ways in which

a process might be carried out; the working biologist can try to verify

one or some of these. He may return saying that something almost works,

and legitimately ask if there are alternative automaton theory ways in

which, within certain limits he wishes to respect, or feels are important,

the process can be carried out.

This makes of the biological automaton theorist, a constructor of

possibilities, and the biologist a verifier of their reality. But of

course the "rules of co-operation" need not be rigid.

[It may have been noted that many of the above problems are remini-

scent of those posed (or at least emphasized and made explicit) in the

last decade in linguistic theory. (Substitute "grammatical" for "viable",

"sentence" or "language" for "organism" or for 'living system" in one's

thinking about biology.) The connections to linguistics lead one to

consider whether other parallels may be worked out. After all, in both

linguistic theory and biological theory, we are attempting to grasp for-

mally some real-world phenomena, to separate it from what surrounds it,

and to describe and explain it, under conditions when only a small finite

part of it will ever be available for our perusal. In both cases there

is the working hypothesis of reasonableness, that there is some way to

explain the phenomena, and that the way is discoverable.

Some other points might be noted. In both linguistics and in biology,

probability and statistics have traditionally (with some exceptions) been
the dominant sorts mathematics employed. When however, recent scientific
linguistic interest shifted from analysis of data to explanation of the nature
of linguistic theory, the discrete mathematics of algebra, graph theory, logic,
recursive function theory, etc., began to be employed liberally. The same
will probably be true of biology; analysis of "large heaps of small pieces of
reality" leads "naturally" to the use of probability and statistics; construc-
tion of explanatory theory for much of biology will lead to the increasing use
of discrete mathematics, and in particular discrete automaton theory (while
probabilistic automaton theory will undoubtedly become a major tool for evo-
lutionary studies). (This analysis-synthesis distinction can also be seen in
cryptology: crypto-systems are attacked by applying probability and statistics
to (if available) large amounts of data; crypto-systems are constructed (and in
the analytic process, reconstructed) employing algebra and logic.)

The connections between linguistic and biological theory can be developed
considerably. Note that as the native speaker produces grammatical sentences
without being able to explain the process, so can organisms produce organisms,
employing their reproductive systems, without being able to explain the process.
This may seem to be a frivolous, even spurious link, but I believe it has some
suggestive content at the very least.

Yet another connection between my picture of biological theory and the
contemporary linguistic thoery lies in the attitude toward the subject matter
of the theory. I wish to set aside the particular physical expression of an
individual organism (which may be faulty) and concentrate on the ideal expression
of an organism (the failure of an organism to achieve its ideal form is of
course very interesting; but I do not choose to deal with the matter at this time).

[Note that by "ideal expression" I do not mean a species archetype.
Variation within a species is extremely important in, for example, studies of

evolution. What I am interested in here is what an organism would become if its growth was not perturbed or distorted by accidents or environmental anomalies.] This notion is meant to be quite analogous to the _competence_ vs. _performance_ distinction so often made by Chomsky. (This is the distinction between the ideal speaker-listener who knows his language perfectly, whose _competence_ is complete, and the actual linguistic performance which may (and usually does) fall short of "perfection").

That is, in this view of biological theory we are concerned with all of the undistorted presentations of the viable. Physical forms of organisms will always deviate from the ideal. (Please note again that what may seem to be Platonism here, does not at all mean that I believe these forms to be members of a fixed class or immutable archetypes or pre-existent in some ideal world.)

How far should the analogy to linguistic theory be pushed? Linguistic theory has been an extremely active field in the past decade, and an immense amount of new knowledge (or at least new ideas and new approaches) is now available. Paradoxically, there has been the noise of an earth-shaking break-through, but disagreement about whether such an event has actually taken place. If biological theory should spend a decade on a similar path, would it be any better off? Linguistics is notoriously subjective. Is biology more concrete, more graspable? Since biology can, easier than linguistics, escape such slip-pery vagaries, will firm results come faster? Can biological theory end up providing methodological guideposts to the linguists?]

After such speculation, how ought we to proceed? What features ought our formalisms to possess? These questions will be the subject of the next section.

6.2 _Some General Criteria for the Formalism_. What "ground-rules" ought we to observe in setting up a formalism for the description of biological processes? What is necessary in a formalism for expressing biological theory?

More specifically we will consider whether it is reasonable to limit ourselves to models which are deterministic and discrete, and whether indefinite growth, and universal computation properties ought to be included or excluded.

In regard to limiting consideration to deterministic models, it perhaps ought first to be said that the biological events of individual organisms we wish to account for lie well short of the sub-atomic world at which observational uncertainty arises. We are therefore not obliged to resort to statistical and probabilistic techniques, although these techniques have been useful in biology and can be expected to continue to be useful.

At the level of physical activity at which biological theory for individual organisms arises, we take it that the events are completely deterministic (though mostly as yet unknown to us, or only partly known). Certainly by the time we get to the cellular level, there is general agreement that we are in the realm of deterministic cause and effect properties.

[Some may argue that this "restriction" to the deterministic implies that life is merely machine-like, and that machines can not grow, reproduce, "think" etc. I can only point out that humans, trees, goldfish, and the dog next door can provide appropriate convincing examples to the contrary; this observation is not, unfortunately, likely to convince the unconvinced.]

We may at times assume chance effects impinging on systems but this is merely determinism in another guise; it is an unwillingness to bring phenomena into the analysis which are otherwise peripheral to the system under consideration, not a disbelief in the essential determinism of the system. (We speak of "random" mutations, not because we think they are essentially not deterministic, but because either with our present knowledge we can not account for them, or to widen our biological system to

include infrequent "intrusions" (eg. cosmic rays) with extremely complex

properties serves only to divert attention from the more biological parts

of such a super-biological system.)

Since we wish to focus our attention on biological theory, we must

be crucially interested in explanation. One feels one understands when

one finally "knows exactly what takes place" "knows exactly how it is

done," "how it works". Presumably even the agents of the Tobacco Institute

would be convinced that there is a connection between smoking and lung

cancer if the exact process from tars and resins to disturbance of cell

function was precisely and completely documented. That the words "only

statistical" possess the power to weaken the conviction of some persons

about the link be-speaks not merely a lack of sophistication on the part

of these persons, but genuinely reveals that mechanistic explanations are

important if we "really" wish to communicate.

Examining next the adequacy of discrete models for biology, it can

at least be said that a great deal of biology is appropriately describable

employing only discrete models. At the molecular level there is a sense

in which it is reasonable to say that chemical bonds are present or not,

that molecules will join or form or not, and that intermediate conditions do

not enter in any biologically meaningful way. That is, there is a level

of description, pertinent to some biological activity of interest, and

this description employs only discrete states and events. Nucleotides

form discrete elements of a code, enzymes are switched on or off, etc.

At the cellular level, cells (or possible cell sub-units) and their

discrete states can be taken as the pertinent entities of the model, and

transient states and conditions and particles are not likely to be germane

to the questions of growth, development, and control we wish to pose.

[While we choose here to emphasize the usefulness of deterministic,

discrete formalisms for biology, it is certainly true that chance and pro-

bability, practically speaking enter into the physical world whenever we

practically attempt to gain knowledge or conduct our affairs. The question

is whether this is relevant to the issues with which we are here concerned.

In a discrete, deterministic model of human growth and development, the

right and left halves of the body might exhibit exact symmetry; it would

seem ridiculous to ban such a model because "in reality" people don't

turn out quite that neatly.]

We come next to the question of allowing for organisms of indefinitely

large size or of indefinite duration of life. Should indefinitely large

numbers of elementary units be permitted to enter into the model of a

particular instance of an organism? Is potentially infinite growth to be

permitted? Since there are no (to our experience) biological organisms

which are a hundred miles long, and no single ones that have been metaboliz-

ing for a million years, shall we, by suitable restrictions in our forma-

tion rules, attempt to prevent such anomalous phenomena being expressible?

[Of course, there has been at least one "organism" a million years

old, and hundreds of miles in extent: the biological system of this planet.

I certainly would not want to exclude this instance of a living system;

but at this point in the discussion organism is being used here in a narrower

sense (in fact the usual sense).]

All organisms we are acquatined with or are likely ever to be acquainted

with (or even "reasonably" likely to conceive of) grow only to some finite

size and thus it might be argued that in any formalism for modelling

biological objects or events only a finite number of elementary units should ever be employed. This may seem to be a sound decision, but we should not impulsively restrict what can be expressed by setting arbitrary limits on size. At any moment, organisms will be finite, but there is no good reason to restrict biological systems to some fixed finite bound. The situation here seems to me quite analogous to what Chomsky has argued in respect to sentences: no sentence a million words long is ever likely to be uttered, but surely a grammar must permit this. There may of course be a "bizarreness reaction" to such a sentence, should it actually be uttered; in a biological system, a kidney a mile across would be even more bizarre, but our system ought to permit organisms with trillions of cells. There is no "natural cut-off point" on the size of organisms (not that there might not, in "realistic" cases, be increasing or even impossible metabolic and structural strain) beyond which we should disallow it. If the biological (or biological-like) system to be modelled is completely open to materials and energy, then what will happen will depend on the cellular control mechanisms. It may be very instructive to assume unlimited resources to support what would in reality be a biological catastrophe. (Cf. the "population explosion".) The real issue is: does the formalism include properties which if properly brought to bear will control growth, permitting both the "normally normal" as well as the "abnormally normal"? The system should allow the expression of all biological phenomena, not just "well-regulated" and aesthetically approved phenomena.

[It should be clear that we wish to allow organisms which may be both subject or not subject to the "laws" of relative weight, volume, etc. (For a discussion of this, see Bonner.)

It probably will be best to concentrate on the "progressive" factors

of growth (the creation, displacement, and differentiation of organic

material) and, for the moment, give less emphasis to the "limiting" factors

of growth (except where the limiting factors are themselves part of the

closed controllable system of the organism). It is of course important

to see how the consequences of progressive factors are modified by differ-

ences in the environment, but if environmental effects can be isolated and

their effects possibly excluded, it is probably wiser to do just that and

so establish the characteristics of less complex systems before working

on the more complex.]

Once we permit the indefinitely large, we open the possibility that

there may be undecidable properties about our formalism and what our

formalism can express. If universal computation creeps in (as it can

when given the slightest toe hold) we shall repeatedly expect to find that

we can not in general tell if processes will halt, whether they will

produce particular patterns, etc. Should such properties be barred by

fiat and the formalism consequently constrained?

I think we should take as a "working article of faith" a sort of

Turing-thesis extended to biological phenomena. (The Turing-thesis asserts

that our intuitive notion of what can be "mechanically calculated" coin-

cides precisely with what can be computed by a Turing machine.) That is,

not only is our intuitive notion of what is possible in a biological system

expressible by a Turing computation, but what is possible for a Turing

machine is likely to be possible for some (theoretical) biological system.

The first half of the above "Turing-biology" thesis (that any biologi-

cal system is a mechanism) is relatively easy to swallow; we accept it as

a working hypothesis every day. The second half (that any Turing-computa-

tion can be carried out by some biological system) is, of course, not as

likely to gain easy assent. We know first of all that "memory" must be indefinitely expandible, but that "real" growth ultimately stops. Let us however, on the grounds given earlier, accept that some organisms and their behavior can not be adequately understood unless we permit indefinite growth. Then, at least for theoretical systems of biology, our thesis becomes a possibility. Under this assumption (and a few other minor idealizations mostly having to do with deterministic accuracy) it is quite easy to present arguments that any Turing computation can be carried out by some biological mechanism. The swiftest solution at this point is merely to offer an idealized human as our biological mechanism. (Our mecano-man would be "more than human" because of his storage capacity, his tireless-ness, and his accuracy; he might be viewed as "less than human" because of the rote nature of his tasks.)

In the earlier parts of this paper we have already exhibited several theoretical biological systems of varying degrees of "believability", which possess universal computation properties. I think such systems will repeatedly arise in theoretical work, and I see no reason deliberately to try to exclude them at the beginning. They may of course not throw much light on biological problems, and that is the crux of the matter. It will be specific (perhaps non-universal) sub-machines of the deterministic realm which will possess the explanatory power we seek here. Thus though the formalism should be capable of expressing universal computation and construction, for any particular organism or class of organisms severe constraints may have to be applied to bring out clearly the "essential" nature of the processes of interest.

6.3. Some Criteria for a Formalism for Growth and Development of an Organism. Suppose that we wish to concentrate on systems for cellular growth and development. What special criteria, conditions should we respect?

Should only a finite, fixed number of kinds of elementary structures

(cell automata, molecular automata) be permitted? Should there be a bound

on the number of states or on the degree of complexity of the elementary

structures?

In general (it is my belief) no arbitrary or fixed bound should be

set on the complexity or the number of kinds of elementary structures or

automata. This is not to say that we may not wish from time to time,

to demonstrate that some important process can be modelled with a few

kinds of simple automata. Von Neumann showed that universal computation

and construction (including self-duplication) was possible employing clusters

of a single kind of (29-state) automaton. Von Neumann did this however, as

part of his campaign to secure conviction on the part of the reader that

self-duplication in a reasonable sense had been demonstrated. To secure

this conviction von Neumann eschewed "super-rules" of the sort which,

for example, take one machine or a system of machines of any degree of

complexity and any sort of form, and, with a single application of the

super-rule produce a copy of the whole machine organism. The offspring

had to be constructed, piece-by-piece, by the parent machine   Achieving

this solution not only demonstrated his design skill, it earned assent

to his major assertion: that machines could duplicate themselves. But

although von Neumann imposed upon himself these strong constraints in his

biological-like system it does not follow that we must always do likewise.

What should be recognized is that permitting automata of arbitrary complexity

may weaken the significance of the results (at least in a mathematical

sense). Showing that Lindenmayer's system is capable of universal computa-

tion is extremely simple (as we have seen) if arbitrary finite automata

and arbitrary growth rules are permitted.

Of course for every particular phenomenon we wish to model we will

usually have a fixed number of finite cell-automata or molecule-automaton types. So although we should not feel we must try to do everything, (model every biological object or process) utilizing a single primitive or set of primitives, most explanatory and informative power will be lost if we repeatedly insert new "science-fiction" devices to handle every trouble-some new situation. Occam's razor will continue to be a useful tool and the god-from-the-machine to be regretted, not welcomed. What primitives we will employ will depend on what we intend doing; and if what we intend doing is modelling the development of a single organism, then it may be desireable to make cells our basic unit and to make each cell genotypically identical to every other cell.

What other characteristics should an automaton system possess? If every cell is not only genotypically identical to every other cell, but is also "windowless", then it is difficult to see how any distinctions of form can develop in the cell system. Information, and consequent alterations of behavior, must come from outside the individual cell, either from other cells, from the environment, or from both. A cell automaton system should undoubtedly allow for information from both other cells and from the environment (although it is of considerable interest to see what can be accomplished in the way of coherent growth and development using no information from the environment, or "very little", or using various kinds of severely restricted communication between cells, etc. What is the least amount of communication of information across cell membranes which will still yield complex structures?)

[S. Ulam and others have experimented with "cell" growth systems with very restricted rules and "reach"--the immediate "neighborhood" of a cell consists of at most directly impinging cells. Very complex appearing patterns have been produced from such very simple rules; questions about

the amounts of information required (and thus the nature of storage and storage sites) to determine growth can be extremely tricky.]

It seems plain that a cell-automaton system should allow the possibility that cells located differently may behave differently, and that their differences, might at times, have to be communicated to other cells.

In several systems (eg. Tag) two antecedent conditions (two distinct active site conditions) which are only "slightly different" in terms of the order or presence of coding components, can have vastly different consequences. Although this does often happen in the biological world (for example at the locus of sex determining conditions) some rules about continuity of effect ought to operate here. If a single element of a genetic code is disturbed it is understandable that (metaphorically) the polar-bear turn out brown, or his toes be unfurred; the error should not convert the bear into a cryptomeria. Closeness of form of result ought to reflect some closeness of form of determining code, and vice versa. The arbitrary assignments of consequences which the free-wheeling world of recursive functions allows (an arbitrary vector, no matter how long, yielding the value true, while another vector differing by one digit in one place only, yielding false) can weaken greatly the biological relevance of many formalisms.

[Natural organisms also often have great resilience. A minor injury (perhaps by definition of "minor") generally has non-lethal consequences. The von Neumann cellular system however, is quite "fragile"; the deletion of a single pulse almost anywhere usually progagates consequences which preclude the successful attainment of any coherent growth or development.

In defense of von Neumann cellular machines, however, it should perhaps be pointed out that their evident fragility may be largely owing to the fact that examples of their activities usually consist of

"stripped down" cases of genetic reading and transcription, and consequent
morphogenesis. The biological counter-parts to the von Neumann self-
reproducing machine would be the reading of the genetic code and the
activities of cell replicative machinery, which perhaps _is_ rather fragile.

With a little consideration of sensory signalling of damage, along
with redundancy techniques, much of the analog to musculature and to life
support activities could be made self-repairing and fairly secure in a
von Neumann cell machine.

There is in the von Neumann cellular system another discontinuity
possible which is however oddly suggestive of some unusual but obviously
reasonable biological behavior. In the von Neumann cellular system,
an original or "ur-parent" machine need not resemble the first-born at
all. The ur-parent, to be sure, must possess certain of the construc-
tional-environmental properties of motherness. Aside from this though,
the form of the mother machine on the one side, and the form of child-
producing tape and the child machine itself on the other, may have little
resemblance. A child, machine or human, can be nurtured in the womb of
a being having little relationship to the genetic code or the form of
the child, and the child can indeed, "leap from the primeval slime", if
it is properly constituted slime. Under the "ground-rules" of the von
Neumann system however (where the genetic construction tape is copied into
the offspring machine) once the series of (false) "duplications" has taken
place, further offspring will be identical.]

It seems also reasonable _not_ to have different laws of "physics
and chemistry" at different points in the "same" system. If a particular
condition at an active site yields one sort of action, then the same
condition at another site should also yield the same action. In design-
ing a large system piecemeal with many subsystems,

these "violations" of the rule of consistent action may be frequent. On

the other hand the differences may be only superficial or may be reconcilable.

This is analogous to the problem of two automata which are "really" the

same, though labelled differently.

It should also be kept in mind that in biological systems there may

exist a "tactic" which would permit the same or only slightly different

codons to have very different consequences. A single strand of messenger

RNA could yield vastly different proteins providing the soluable RNA was

different. That is, there couldbe several sRNA fragments, each matching

at one end a particular mRNA codon, but having different amino acid attach-

ments at the other end (though this has not been observed to be the case).

6.4. Some Comment on Universality. It may be useful at this point to

discuss further what is meant by universal computation and universal con-

struction. When we say that a device (or a class of devices) is capable

of universal computation, we mean that the device or class can compute

the same class of functions as any known universal computation class, as,

for example, the class of all Turing machines. One way to show that a new

system has universal computing capability is to show that for any given Turing

machine whatsoever, the new system could perform the same computations. This

does not mean that the new system must synchronously behave, step-for-step,

like any given Turing machine. Indeed, it is often the case that several

operations in the new system are required to carry out a single Turing machine
operation.

If the repertoire of Turing machine operations (as well as any sequential

combination of these operations) can be carried out by the new system,

then the new system can compute what any Turing machine can compute, and

thus itself possesses universal computing power. Since there are particular

single instances of Turing machines which are universal, an alternative

way of showing that a new system possesses universal computing powers is to

show that the activities of any of these single universal Turing machines can be carried out in the new system.

In this latter case we have explicitly a single machine which possesses universal computation properties. This machine can compute anything computable, but again it does not necessarily directly carry out its operations in precisely the same manner and rate as any other Turing machine. The question whether there are reasonable stronger senses of universal naturally arises. Is there a machine (or machine system, or machine descriptive formalism) broad enough so that for any given machine not only is there a machine in the new system which can carry out the same function, but in addition will carry it out in a manner "the same as" that of the original machine.

The crux of any further discussion of this issue lies of course in what we might mean by "the same as". The precise characteristics of "sameness" will be important relative to what we wish to do, how we intend using the systems, what we are interested in accomplishing. A similar issue arises in regard to universal construction.

The von Neumann cellular system permits both computation and construction (indeed it may be unprofitable to try to distinguish them). Machines capable of universal computation can be expressed in the von Neumann system. Thus, the computation of any Turing machine can be carried out by some von Neumann cellular machine. This does not mean however, that a von Neumann machine exists which can carry out the computation of any given Turing machine in precisely "the same" manner as the given machine. Thus immediately we see that there are limitations on what can be constructed by machine, in the von Neumann cellular system.

There is a von Neumann machine which is a universal constructor. What then do we here mean by universal? We mean that there is a particular

von Neumann machine which can carry out the construction of any other von Neumann machine which can be constructed. (Among the constructable machines is the universal constructor itself, so self-reproduction is possible.) Among the von Neumann machines which can be constructed are machines which can carry out any computation; a universal computer can be constructed by the universal constructor.

The following important limitations should be noted. First, there are machines which can be placed in the space from the outside but which can not be constructed from within the space. That is, the von Neumann space will permit the existence of certain machines but not permit their construction by any machine itself operating within the space, including construction by the universal constructor. Secondly, there are many other machine behaviors which are not expressible, with complete fidelity, in the space at all.

[One place in which this limitation reveals itself is in the fixed time delays between cell-automata; because of this, certain easily describable pure switching activities are not expressible directly since to achieve the result in a von Neumann space may require some considerable amount of time, while in some switching formalisms the action is considered to take place "instantaneously". Mukhopadhyay (employing results of Mc Naughton) has shown however that, at the cost of considerable structural modification, any finite automaton input-output behaviors can be achieved at the cost of some time displacement.]

This second limitation strikes directly at an issue critically important to us if we wish to design a formalism for the description of any biological-like processes. Any particular von Neumann space may be unable to express (in terms of an embedded machine) some biological process of interest. John Holland has attacked this difficulty frontally by attempting

the design of a system of "universal spaces" permitting the embedding of any machine-like processes whatsoever. Since we do not wish, at this point, to adopt his clearly sensible strategic approach, but rather to employ less than "completely universal" systems, some justification of our approach is in order.

If we employ formal descriptive systems which are less than "completely universal" we can expect that many biological processes of interest may suffer great "structural" distortion if expressed in the formal system. This great distortion may seriously vitiate the relevance of any results obtained in the formalism for the original biological context. This is certainly a disadvantage. On the other hand, one can start with the less-than-completely-universal formalism and develop its properties, keeping in mind possible applications to biology. While not as all-encompassing a strategy as beginning with a completely universal formalism, it is a sound and reasonable research tactic; we shall repeatedly employ this tactic in what follows.

APPENDIX

(To Section Six)

Here is yet another universal biological-like system. It has the
(possibly slight) advantage over most of the earlier systems that its
"operating framework" is less rigid and more biological; I have tried
to make it as "wet" as possible.

In order that the system work properly, the following must be
granted: operations must be carried out deterministically and in the
proper sequence and temporal relationship, unlimited amounts of food,
energy, and time are available. More specifically, the system consists
of two "bags" of cells; the bags may be required to expand to hold
indefinitely large but finitely many cells. The action of this system
consists fundamentally in doubling the contents of one bag and halving
the contents of the other (or possibly the contents of both bags remaining
the same); at this doubling, halving, or remaining the same, the cell-
type in the bags may change; that is, the offspring cell after one of the
above acts may differ perceptibly from their parents (we can require
that the parent cell generation die after producing the necessary progeny,
so that each generation of cell populations consists entirely of newly
formed cells).

The doubling and halving must be done accurately. In the case of
the halving operation, we must be able to detect whether there will be
an odd cell left over. In the case of the doubling operation, we must
be able to add a single cell to that bag (beyond the doubling which will
have taken place). Whether a bag of cells will be doubled in number,
halved, or remain the same, will be in part, a function of the type of
cell in the bag.

This completes the basic requirements. We now further describe how the system works.

Let us assume we already have some number of cells in each bag. The number of cells in one bag is to double and the other to halve. The cell types themselves in each bag will, in the final analysis, determine this. (We could, by the way, use a single bag of cells providing we allow sufficient distinction of cell types so that we could distinguish and work separately with such a mixed cell population; for clarity we employ the two bag concept.) The newly generated cells may be (as we have noted above) different from their parents. The original cells may be required to produce offspring distinguishably different from themselves; the parent cells will die, lyse, or otherwise be eliminated.

Thus, for doubling, we could first induce cell pairing, then the production of four offspring per pair, then the death of the parent cells. Should there be an odd cell unpaired, we must require that it, alone, produce two offspring, and then die. As a function of the kind of parent cell, we may have a single extra cell of the new type, produced and added to the population.

In the case of halving we must first know whether an odd cell will remain (for the nature of future offspring may depend on this). We might picture this as happening in the following fashion: the population of cells which is to be halved first pairs; if there is a cell left unpaired, this environmental fact results in (say) a "chemical message" to that effect. The population now halves (say by each pair producing a single offspring and the parent cells' deaths). The odd cell produces a single cell of the new kind and the original odd cell dies.

This completes an informal description of a cell system which possesses

universal computational properties. None of what we have asked of our

system is, separately, "un-natural" or "un-biological" (aside perhaps
from some$_\Lambda$precision of behavior); each of the separate processes could

unusual degree of

be exhibited in some real biological system. The conjunction and opera-

tion of all these processes in the required manner may very well strain

the biological credulity.

As for the universal computational properties of this system, let

us make them more apparent. There is known to be a universal Turing

machine having 50 states and 2 tape symbols (result of Shannon). The

tape configuration of any of such a class of Turing machines is completely

described by the pattern of zeroes and ones (say) appearing to the right

and to the left of the read-head, and by the zero or one presently under

scan. If we add to this the internal state of the read-head, we have a

complete instantaneous configuration description of the machine. If for

every internal state/scanned-symbol pair we have a (unique) number (which

we can have, since the number of both states and symbols is finite, and

indeed they need be no larger than 50 and 2 respectively (for a total of

100 combinations)), and we view the pattern of zeros and ones to the right,

and to the left, of the scanned square as two (binary) numerals (with the

least significant places closest to the read-head (i.e., in the conventional

fashion on the left of the head, and in reverse fashion on the right of

the head)) then the complete instantaneous description of the machine

can be given by three numbers: one for the tape pattern to the left,

one for the tape pattern to the right, and one for the state/symbol pair

of the read-head. The tape numbers may get indefinitely large, the

state/symbol number is, as we have seen, finite and fixed for each machine.

Let us conceive of our read-head as being able, as a consequence of

its state/symbol pair, to order and accomplish the doubling and halving
of the numbers which represent the binary patterns, right and left, on
the tape. It will be seen that a move of the head to the right means
halving the right number and doubling the left (possibly with adding
one to the left should there be a 1 under the read-head before the shift
right). What the new square the read-head rests on (the square to the right)
will depend on whether the right number was odd or even (the content
of the low-order place); the next action will be dependent on this odd
or even-ness.

The significance of the halving or doubling of a cell population,
and the problems of the odd cells, should now be apparent. We must now
show how we can get rid of the read-head itself. Once we do this, we
will have the two cell-bag biological-like system we described above.

We can eliminate the read-head by transferring much of its control
function to the cells themselves. The population that is to be halved
will, as a consequence of its own type and the presence or not of the odd
cell, determine the appropriate next cell type. The halving population
may have to provide a (chemical) message to the doubling population so
that it will produce offspring cells which correctly double or halve at
the next generation.

If a cell type produces no further action, or if the system "cycles"
by producing the same number of the same types, then the process has
"halted." For every 2 symbol Turing machine there is a cell system of
the sort here described such that the machine will halt if the system
halts. Since Shannon has shown that there is a universal two symbol
machine, the class of two symbol machines is universal, as is our cell
system.

(Since there are only 2 x 50 = 100 symbol/state combinations, our cell system will require at most 100 different cell-types (although more may be required to take care of some odd-even cell-type cases, and some problems of chemical stimulus, and the like.))

(This doubling-halving universal cell system, with the read-head eliminated, is largely a translation of ideas of Cocke and Minsky used in showing the universality of a "small" Tag system.)

Example

$m = 13$ $n = 9$

| | | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | | | |

S_i
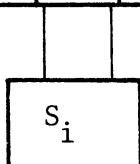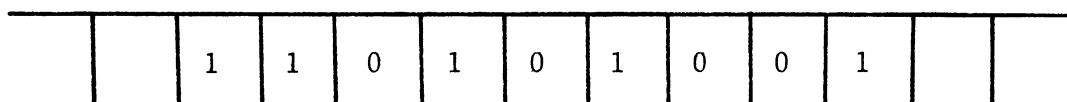
Description
$$\begin{cases} S_i(0) \\ m = 13 \\ n = 9 \end{cases}$$

Below: Result of move to the right; no tape change.

$m = 26$ $n = 4$

| | | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | | | |

S_j
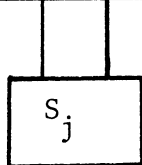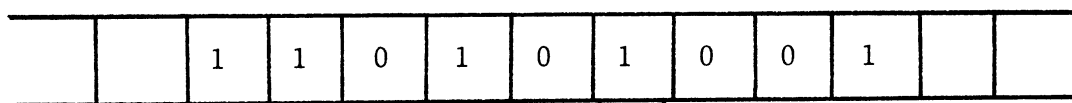
Description
$$\begin{cases} S_j(1) \\ m = 26 \\ n = 4 \end{cases}$$

Turing Descriptions Given by 3 Integers.

Simulating a Right Move (no print) with Control Scanning 0.

Bag L   Bag R

| m = 13 cells | | n = 9 cells |

Control
System

$S_i(0)$

"double" signal   "halve" signal
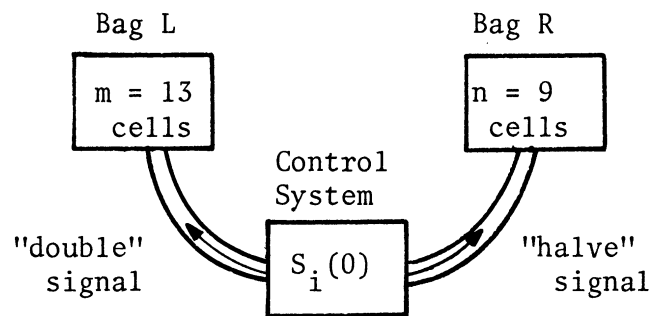
"Status" $S_i(0)$ results in sending a "double" population chemical signal to Bag L, and a "halve" population signal to Bag R.

The intermediate result is shown below.

Completion of Initial Phase (and Start of Final)

Bag L   Bag R

| m = 26 | | n = 4 (+1 old cell) |

$S_i(0)$
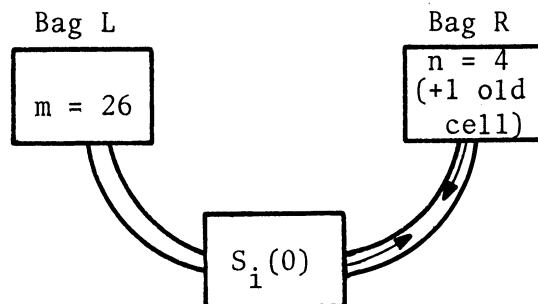
Status must now check for presence of single old cell. Upon receipt ("destructive read-out") of information on presence of excess single old cell, the following description completing the cycle results.

Bag L   Bag R

| m = 26 | | n = 4 |

$S_j(1)$
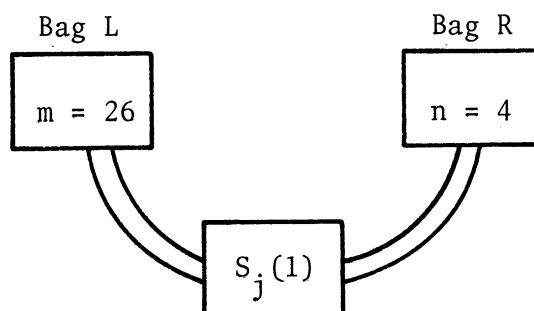
Completion of Final Phase.

Simulating a Right Move (no print) with
Control Scanning 1 (or printing 1).

Bag L   Bag R

m = 13 cells   n = 9 cells

"double" signal ← $S_i(1)$ → "halve signal"

Initiation of Process.

Bag L   Bag R

m = 26   n = 4 (+ 1 old cell)

1 duplicate signal $S_i(1)$ "echo" read-out
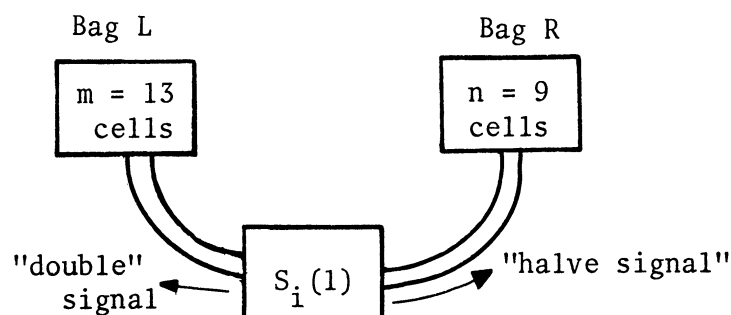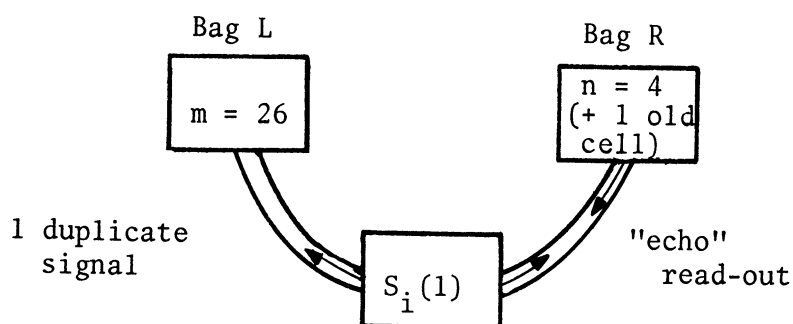
Completion of Initial Phase and Start of Final Phase

Control Status checks as before for excess single cell in Bag R; in addition control sends a "chemical" stimulus to Bag L causing a single cell to duplicate (increasing contents of L by one).

Bag L   Bag R

m = 27   n = 4

$S_j(1)$
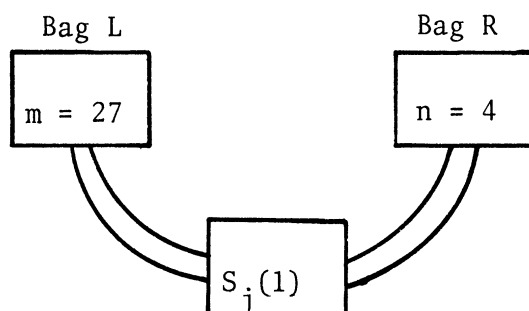
Completion of Final Phase.
System ready for next action.

So the proper behaviors (statuses) can be attained by doubling and halving, and by "single duplication" signals, and "destructive read-out" of presence of single excess cell. (The control-endocrine system can accomplish this by sending a "lyse" "signal" specific for the earlier cell type, and then monitor the cell bag for lysis products-debris.)

Now let us see to what extent we can let the cells themselves take over the control system functions.

We begin by increasing the kinds of cells, introducing for every earlier cell type, a type which combines these earlier properties (doubling, halving, indica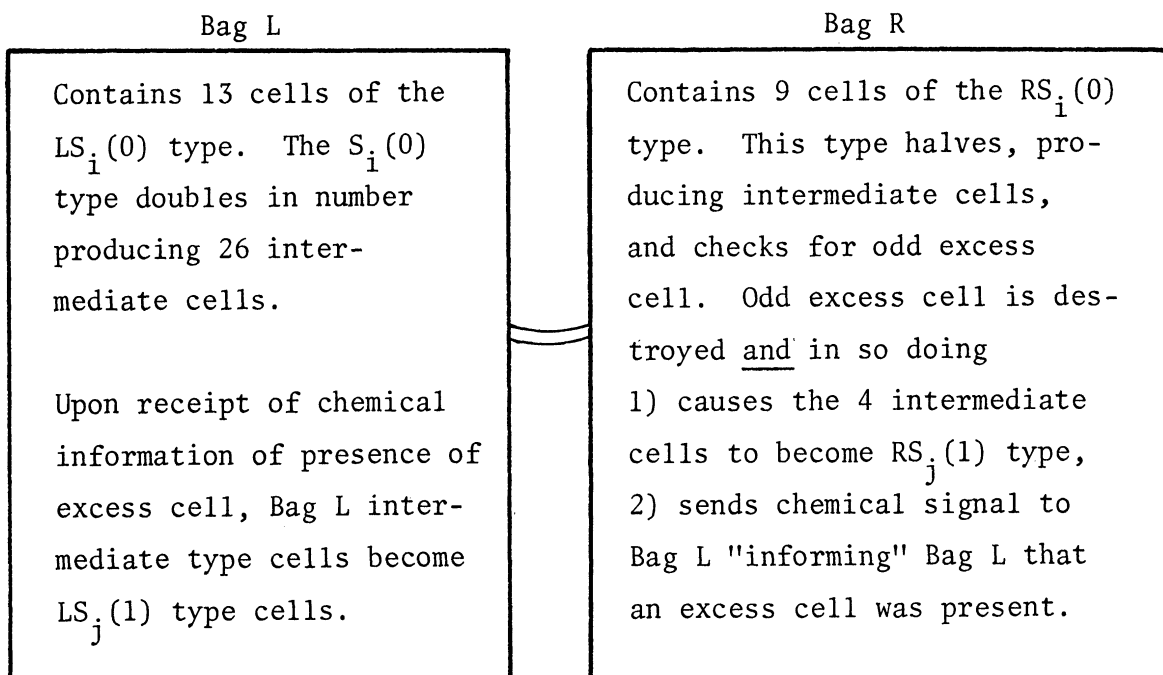ting single cell excess, single duplication) with control status properties, so that control of the "computation" will take place locally, in the cell bag itself, and not be directed from an "outside" unit.

<div align="center">
Simulating a Right Move (No Print) with (Cell Bag)
Local Control (Same Computation Example).
</div>

Bag L	Bag R
Contains 13 cells of the $LS_i(0)$ type. The $S_i(0)$ type doubles in number producing 26 intermediate cells. Upon receipt of chemical information of presence of excess cell, Bag L intermediate type cells become $LS_j(1)$ type cells.	Contains 9 cells of the $RS_i(0)$ type. This type halves, producing intermediate cells, and checks for odd excess cell. Odd excess cell is destroyed and in so doing 1) causes the 4 intermediate cells to become $RS_j(1)$ type, 2) sends chemical signal to Bag L "informing" Bag L that an excess cell was present.

It is clear, without continuing the description further, that all the

desired "computational" actions can be expressed in this format.

I have been assured by a cell physiologist that none of the separate biological systems actions here required is impossible; he (quite right-fully, I suppose) refused to discuss the probability of their conjunction in a single system. And there is at least one other property which must be embodied, to make the system believable: with control vested in the local cells and their conditions, the communication between the cell con-tainers should include some set of synchronizing start-finish signals, or else one cell bag could race out of phase with the other.

The final step in this process, is to put everything inside the same membrane envelope. As long as right and left cell types remain distinct, and interact only in the desired fashion, we can have any Turing machine computation being carried out by a bag of cells. Recalling results for Turing machines, we will not be able to tell in general of such cell-bag systems, whether they will cease growth or continue to inter-act forever. We will also be unable to tell whether for all numbers and types of R and L cells, whether any particular number and type will ever occur.

(It should be remarked that there is nothing special about doubling and halving the cell populations; tripling and dividing by three might, for example, have been used. It all depends on the underlying number base which in turn depends on the cardinality of the tape symbol set of the Turing machine imitated. Doubling and halving was chosen here because it could most readily be made biologically "reasonable".)

(Note also that the cell-bag system described here is vulnerable to the criticism that it can attain such a size that communication of information to cells and from them can intolerably slow the computation.)

7. A LINEAR CELL-AUTOMATON SYSTEM

In Section 4, we discussed how formal biological (or biological-like)

systems could be based on the monogenic normal systems Tag and Lag. In

Section 5 we discussed various formal systems which have been purposely

designed to capture aspects of biological systems. Here in Section 7,

we incorporate many of the notions already presented earlier and describe

a linear cell-automaton system.

[It is difficult to say precisely where the ideas, here to be embodied,

arose. Similar ideas are to be found in Arbib; in a sense all these ideas

have been "blowing in the wind" for years and it would be madness to try

to assign the source; some may very well be original to me.]

7.1. Description. The procedure in presenting the linear automaton

system will be to describe a simple version and then add features, while

commenting, to produce a more elaborate system.

The procedure will thus continue to be exploratory; features will

be added or modified or deleted according as it seems fruitful or interesting.

In the "vine" system of Section 4.2, we did not explain in any detail

how the genetic effect from the root end cells produces the appropriate

action at the growing tip end of the plant. In this section we give such

a more detailed explanation. The underlying system will be Tag (as in

Section 4.1) rather than the Lag system (as in Section 4.2) although a

similar explanation could be given for Lag.

First, to recapitulate Tag system action: we begin with a string of

cell automata; there is a "root-end" and a growing tip end of the string,

and new cell-automata will appear only at the growing tip end. The kind

of growth occurring at each growth step will be completely determined by

a particular root-end cell-automaton. This cell-automaton will send

genetic-growth instructions down the string of cell-automata to the growing tip. After growth has occurred, the root-end cell-automaton which determined the growth will relinquish its control role to a new cell-automaton lying a finite fixed number of cells away in the direction of the growth tip. The next growth phase is determined by this new control cell-automaton. This process continues until (1) a cell-automaton is encountered which has no growth instruction (and does not delegate control of growth to a new cell-automaton) or (2) there is no cell available (the string being too short) to which control can be relinquished.

What properties must our cell system possess in order to carry out actions of the sort outlined above? The growth cell-automaton must be able to send growth signals along the intermediate cell-automata to the growing tip automaton. The growing tip cell-automaton must be able to construct the new cells (or must construct the first of them, allowing the first new one to construct the second new one, etc.); the growth cell-automaton must at the conclusion of transmission of growing tip-information, "delete" itself as well as a fixed number of "next" cells from participation in further activity (that is, a signal must be sent to a cell a fixed distance along the string, that it is to determine the next growth phase, and none of the intervening cells is to interfere with the process.) (Note that the deletion process merely bars a cell from taking part in the present active process; it is not barred from a present "scaffolding" role or some future active role.)

There are many specific solutions to these problems: here is one.

There will be a finite number of kinds of cell-automata. Every cell will be identical in that each will be able to produce any of the other kinds of cell automata. Each cell-automaton will be able to tell when it

is the growing tip automaton; each cell will be capable of producing

one of any of the cell-automaton types. Signals of various types (which

may be viewed as chemical or electrical) can be initiated, propagated,

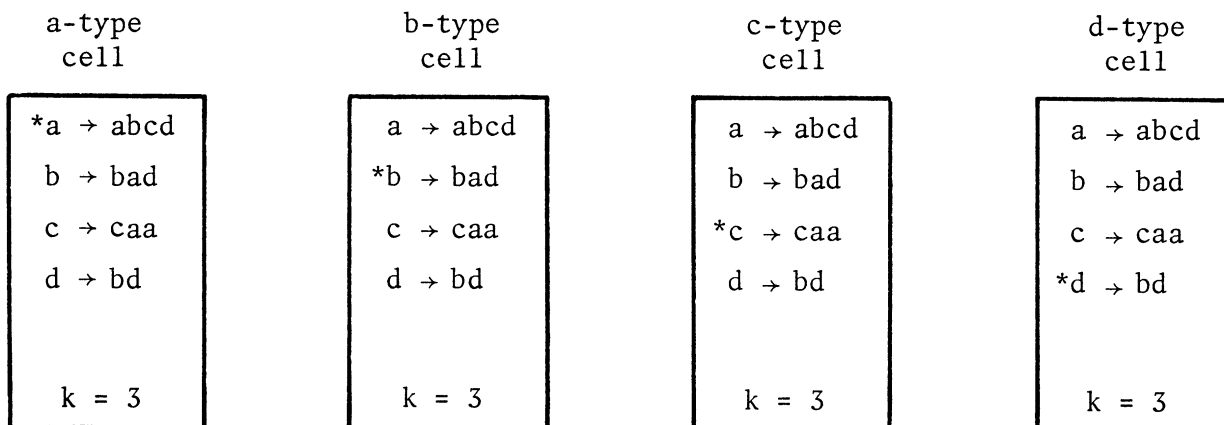and acted upon by each of the cell automata.

We now describe the action of the system in more detail.

We begin with a cell-automaton being designated the growth determiner.

The cell-automaton then goes into a sub-routine unique to its type. This

sub-routine consists in sending sequentially, toward the grow-tip end,

signals specifying each of the new cells to be created at the end. These

signals are sent, each signal separated by a pause, to the next cell-

automaton. Any cell which has another cell on its grow-tip side acts as

a straight-through transmitting unit. The one cell (the tip cell) which

has no cell as a grow-tip side neighbor is in a special state such that

when it receives a grow signal, it produces the specified cell. This

growth takes one time-step (hence the one unit pause between grow-cell

signals). Each successive grow tip cell will produce a new specified

tip cell until the signals cease. (See Figure 7.1 for an example.)

[An alternative method of handling growth would be to have the cell

which is at the tip retain control of growth through the growth phase,

creating new cells to its right, say, pushing the tip farther yet to the

right. In this system, the new cell types all "created" by the original

tip cell, will end up in reverse order (relative to the first growth

method) in the cell-automaton sequence. This convention can be accommodated

to. In this system, we will have lost the idea of "natural" transfer of

growth to the right-most cell, and we must therefore introduce a mechanism

for the transfer of growth to a new spot. Although this complicates the

LINEAR CELL AUTOMATON

Example

a-type cell	b-type cell	c-type cell	d-type cell
*a → abcd b → bad c → caa d → bd k = 3	a → abcd *b → bad c → caa d → bd k = 3	a → abcd b → bad *c → caa d → bd k = 3	a → abcd b → bad c → caa *d → bd k = 3

Initial cells situation (growth control cell designated by ☐).

1. ↓
   [a]

2. [a] [a] [b] [c]↓ [d]

3. [a] [a] [b] [c] [d] [c]↓ [a] [a]

4. [a] [a] [b] [c] [d] [c] [a] [a] [c]↓ [a] [a]
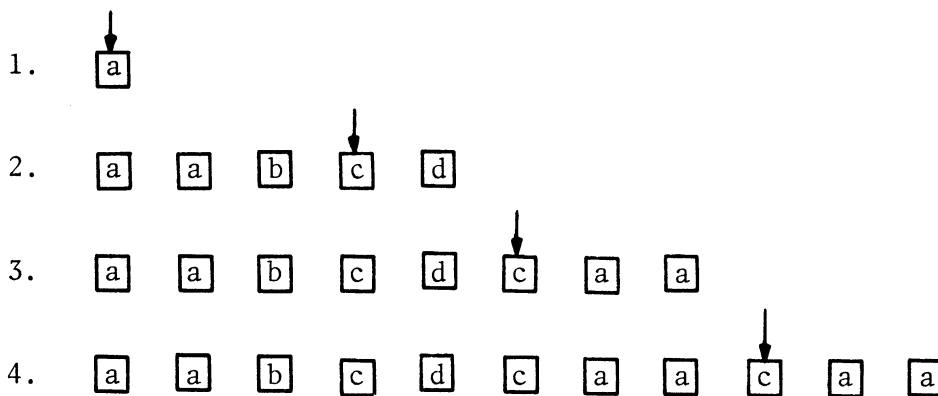
Continues to Grow but Repeats Hereafter.

Fig. 7.1.

growth rules, it also creates new possibilities; in particular it introduces

a mechanism for creating diverse growth sites, a feature which may be

very useful.]

The sub-routine within the growth-determining cell will at the conclusion

of its transmission of initiating growth signals, send a series of "deleting"

pulses. These pulses are intended to remove the next k cells from the

growth initiating process, and in effect designate (or activate) a new

"first" cell, which will then determine the next growth. These actions can

be accomplished in various ways. One way is this: at the conclusion of the

transmission of the growth determining signals, the sub-routine sends k-1

deleting pulses. After the kth pulse the cell turns <u>itself</u> off (the kth

cell); the cell adjacent to the active cell, upon receipt of the first

delete pulse, employs this pulse to go into a state in which it will here-

after act only as a transmitting cell; the original delete pulse is not

passed on, but all successive delete pulses are. Whenever a cell finds that

its root-end neighbor has been in a "turned-off" state for more than one

unit of time, it "recognizes" that it has been designated to determine the

next growth phase.

Obviously no new growth will occur if by turning off k cells, none

are left to initiate new growth, or if a designated cell has no growth

sub-routine.

[Cell-Tag, as so far described requires a means of transferring

control k-1 cells toward the growing-tip end. Above we have shown how this

may be accomplished (employing a series of special pulses). If however

we take advantage of one of our original assumptions about cells and

molecules, namely that they may have characteristics determining the

structural forms they may take, we can eliminate special signalling of

control transfer, and make the structural forces assist in the determining

of the next control cell.

Let us suppose that the "bonding-adhesion" forces cause the molecule-

cells to form a close-packed spiral. If each turn in this spiral is k

units long, then a transfer of control across one set of cell membranes

in the axial direction will give us the desired transfer result without

resorting to special pulses.
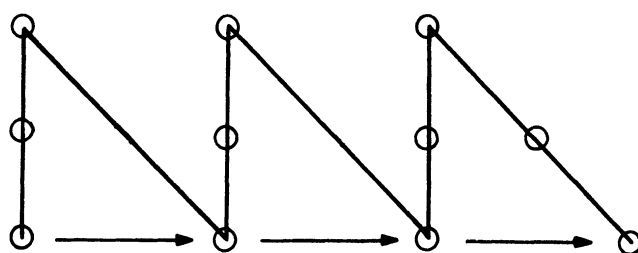
Control transfer.

Fig. 7.2.

In effect many regular patterns of "clumping" or folding of newly

produced cells could bring the sequence of present and next control cells

into juxtaposition. It is interesting to contemplate what can be effected

when the transfer of control is determined by e.g. some spatial contiguity

property that does not yield a fixed k in the linear direction.]

This completes the description of one way in which Tag may be carried

out in a cell-automaton system. Since Tag is universal, so also is this

cell-automaton Tag; since Tag halting is undecidable, so is this cell-

automaton Tag; since the word problem for Tag is undecidable, so is it

for cell-automaton Tag (i.e., we can not tell if any arbitrary pattern of

cell types will be produced). Information is sent in one direction only

in Tag and in cell-automaton Tag (this fact can be employed to show the

universality of Turing machine action on a tape which is infinite in only

one direction).

Minsky and Cocke have shown that there is a universal form of Tag which deletes exactly two letters (turns off two cells); the under-lying universal computation which is mirrored by this Tag system, is that of a two-symbol Turing machine; for every state/symbol pair of the under-lying machine, there is a letter in the Tag system (cell-type). Shannon has shown that there is a universal Turing machine with two tape symbols and 50 states. Given that some additional "punctuation" letters may be required in the Tag system, for universal cell-automaton Tag a few hundred cell types may be required. Notice though that although these several hundred cell-types must be different in some identifiable sense (in the specific growth sub-routine they will employ) they are otherwise identical. Thus, the cell-types can all be similar, each of them having the same information control characteristics, and each possessing all of the sub-routines; they will differ in that each type will have a different pre-determined growth sub-routine. Whenever a cell produces a new cell, it will in effect reproduce itself but will "check" or "designate" the specified growth sub-routine which will be characteristic of the new cell.

7.2. Discussion. Thus with a single kind of cell-automaton, having approximately 200 states, we can perform any (Turing-computable) calculation. If for instance the "calculation" is the construction of a particular string (computation of a particular function) we begin with a single _active control_ cell placed in the environment of (at the left most position of) a string of cells which embody the function to be carried out as well as the argument upon which the function is to act. Notice that active control is then shifted right, through the string: the "function fragment" and its argument are, step by step with the control, interspersed in the cell-states

of the original string.

We have here just expressed cell-automaton Tag as a series of Tag actions imposed by a control cell acting in a cell-string environment. These initial conditions can be pictured as arising in different ways.

1) (As in our example) the initial Tag string and an active control cell are assumed given. ...

2) The initial string may be created by a series of growth impulses originating from outside the immediate cell-Tag environment. That is, a series of grow-pulses dictating the construction of the cell-members of the string are received by the initial cell (which would not necessarily be in active control mode to begin with), growth proceeding in the manner of the description of Section 7.1 at the conclusion of the outside stimulus; at least one cell automaton should then be placed in active control, either by specific direction, or by some internal rule of the system (such as we have suggested in Section 7.1 where the left most non-"turned-off" cell is designated as being in active control.

3) The initial cell string could be created by a special series of "genetic" instructions, perhaps present in every cell. Thus, a single stimulus to a single cell would result in the growth sequence for a string of cells which would then be acted upon, as well as act.

All three of these cases are interesting for their biological analogies, and aspects of all might be employed. In the last case, for example, one or more cells might carry instructions for producing the more complex cell conditions which will (perhaps with some external environmental stimuli) eventually result in the "complete organism".

The flexibility of the model might be increased by incorporating some additional features.

As the system has been described, we have identical cell-automata, except that they differ in their characteristic cell-growth instructions. (They all behave the same way as afar as far as transmission and reception of information signals, cell-construction procedures, active-control characteristics, etc.) We have treated their characteristic cell-growth instructions (the logical consequent of the underlying Tag system rule) as a "wired-in" property. We could however incorporate in our system the idea of a change in characteristic growth instruction. That is, in certain cell environments, under external stimulus, etc. the growth signals disseminated by a cell automaton might be altered. This new property, along with multiple active control cells, and multi-directional propagation of signals, would greatly increase the flexibility of the model.

Such broadening of the system properties raises many issues. There may prove to be difficult "priority" and precedence problems, problems of resolving conflicting instructions, etc.

8.   REPLICATION IN A LINEAR CELL SYSTEM

In this section, a discussion of replication (or self-duplication) of the cell-construction sort (that is, in the fashion of machine self-duplication by separate signals for cell creation at a distance) as employed by von Neumann rather than by local cell fission (which is undoubtedly more "natural") is presented.

There are several reasons for considering this sort of replication in a linear cell system. First, it is interesting in its own right and as an exploration of what can be done; for example the question of self-duplication in a linear cell system might show how morphogenetic processes could take place. In addition attempts to achieve replication will serve to bring out the characteristics, the difficulties, the techniques applicable to linear cell systems. Secondly, it may be of some assistance attacking problems originally posed in the von Neumann cellular system; among these questions are 1) what is the minimum complexity required in a VNCS that it be self-duplicating 2) can self-duplication be detected when it has taken place (i.e., is the set of self-duplicating VNCS recursively enumerable?)

[These problems will be discussed in some detail in Section 9.]

8.1   Von Neumann System Self Duplication

John von Neumann showed how self-duplication can be achieved by a machine; the machine was expressed in terms of cells in a two-dimensional array. In his self-duplication system, the parent constructs an offspring copy of itself, cell-by-cell. In particular, von Neumann did not employ any "super-rule" (of the sort where by fiat, the whole complex of cells which make up the parent machine are duplicated in a single construction act).

While von Neumann's system does clearly possess self-duplication properties, the methods he employs to achieve this are not necessarily very "biological" and (as we have noted earlier, viz. Section 6.3) much of the character of the procedures which he did employ was a product of the particular task he undertook to convince skeptics that machine self-duplication was possible and could be closely and rigorously described. The argument has been won now, and for this reason we shall not feel bound to imitate von Neumann's methods closely; instead we will use his ideas as "jumping-off" places.

It should first be noted that the linear cell system we have introduced is in some respects simpler and in other respects more complex than von Neumann's; for this reason some properties may be harder to exhibit, and others easier. Von Neumann restricts himself to automata having at most 29-states; we permit our cell-automata to have many more states, and thus to possess some rather complex properties. For example, the pulsing, decoding, etc. functions that our cells must perform are carried out, for us, as cell rather than machine functions; for von Neumann these were functions for which submachine units, composed of cells, had to be designed. What are sub-cell activities for us, are submachine activities for von Neumann. In effect we are largely ignoring (at least at this moment) the problem of the precise construction and nature of the sub-cellular functions we require, merely assuring ourselves that these functions can be carried out by finite state devices (and sparing ourselves a great deal of logical design which "everyone knows can be done" and which would probably be largely irrelevant to the biological problems of interest). The result is that the complexity of each of the cells of this linear cell system begins to approximate that of complete constructing units of the von Neumann system.

Making this change means that many broad results can be obtained quite easily.

[An interesting exercise, but one I shall not pursue at this moment is to construct, out of von Neumann cell automata (or similarly simple automata) a cell of this linear system.]

The linear system we envisage here is in several respects _more_ restrictive than the von Neumann system. The most obvious restriction is the use of but a single dimension; because of this restriction to one dimension the straight-forward two-dimensional imaging of many machines or processes is barred. But the restriction to one dimension is not as severe as it first may seem (at least in respect to comparisons with the von Neumann system); the linear process, when compared to the von Neumann system, is more like a broad (though fixed) band of von Neumann cell array.

[This observation will be exploited later in order to develop some problems of self-duplication. H. R. Smith has recently employed the (widely known) concept of expressing any Turing machine behavior by strips of specially designed cells (similar to what was described in Section 5.2). Among his numerous results for this sort of tessellation system: there exists a universal one-dimensional (single strip) cell system in which any cell's actions are (at any instant) determined only by its own state and that of at most two contiguous neighbors on one side, and three on the other, and in which each cell has at most 7 states (a seven state, six neighbor (including self) system).]

The added properties of the individual cells of the linear system mitigates considerably the effect of this restriction.

A rather more severe restriction is that (as we have so far discussed the linear system) information is mainly transmitted in one direction only.

In addition, the rules of the underlying normal system require constant
addition of cells from one end, and deletion from the other; continuity
of the system is preserved only by constant recopying.

[This "sweeping" of active site along the length of the system is
not however necessarily unbiological.

Notice also that in a linear cell system, as so far described, there
is essentially serial activity, while in the von Neumann system branching
leading to multiple parallel activity is possible. (As a matter of fact,
though, the parallel action features of the von Neumann cellular system
have never been much exploited; parallel action has been employed in the
design and construction of sub-machines, but the general view point (pro-
bably owing to practical matters of ease of conception and explication)
has been of serial operation, only one thing at a time going on.) Problems
of extensive parallelism have scarcely been touched, some work on distri-
buted control in iterative circuit computers and the "firing squad synchroni-
zation" research being among the scanty contributions.]

We shall feel free however to alter these restrictions in the linear
cell system, first by means similar to that of von Neumann in his two-
dimensional system, later by various means which seem "biologically more
natural".

In the von Neumann cellular system, a tape is fed to a universal
constructing unit; the tape consists of the instructions for the construction
of another machine. By carrying out the instructions on the tape, a new
machine is produced and set in motion. If self-duplication is our goal,
two copies of the original instructions are employed. Each copy is the
instructions for the construction of the very machine which is operating

on the tape. The constructing machine reads the first copy of the instructions; these instructions result in the construction of a copy of the original, minus however the original tape contents [which have (in part) been (destructively) read and acted upon]. One copy of the original tape will have been preserved however, and the final instructions on the tape which is being read are to copy, twice, the contents of the preserved tape onto the tape of the newly constructed machine. The new machine is now a complete duplicate.

As we have described it, the sole function of a self-duplicating machine is self-duplication. We may however wish the machine to produce some additional "side-line" work which we can view as its useful contribution to the environment, its raison d'etre aside from perpetuation of kind. The instructions for this useful work can be included on the tape, and can be copied onto the instruction tapes of succeeding machines. The "side-line" tasks machines perform may be one feature we can employ in distinguishing among machines.

How might self-duplication (or "replication") take place in the linear cell system under discussion?

Instead of an initial machine composed of a two-dimensional array of cells, we begin with an initial string of cells. The left most cell is "on" and determines a new cell growth (which, as in the von Neumann system can be viewed as either the beginning of a computation or a construction-- the two are treated identically). The initial string of cells (as well as the initial two-dimensional array of cells of a von Neumann system) can be viewed as embodying an argument along with a function to be applied to the argument, or as growth instructions for a machine, and the constructor which is to follow those instructions.

The new cell growth occurs at the right-most end of the string. The location of the control (which is at the far left) then shifts rightward. Eventually all of the "initial axiom" or "argument" or "initial tape configuration" will have been acted upon; the process then continues, "feeding" on newly produced secondary rather than the original or "primary" string of cells.

If in the course of rightward growth, an exact duplication of initial conditions is reached (in the sense that at the conclusion of a growth phase the initial string of cell types re-occurs, and the leftmost cell of the new "initial" string is in control) then we can say that replication has taken place. (We might include some "detach" command which will physically separate parent and offspring.) It is also clear that, the system being monogenic, this duplication will repeatedly and regularly recur.

Can this replication take place for an arbitrary string? It can, providing the system of cells is sophisticated enough always to allow certain copying and growing routines. What must our cell system be able to do to replicate? What procedures can be followed?

[Some linear cell systems are not "sophisticated" enough to possess universal construction properties. The problem of the minimal system for universal construction (including self-duplication) is very intriguing. If systematically attacked, something useful may be revealed; on the other hand, this sort of activity may evolve into more grotesque "low score" competition of the universal computation sort.]

Let us now discuss in somewhat more detail the replication problem for linear cell systems. Since the original cell string will be deleted in the course of being read and acted upon, we must somehow preserve a copy of it. This can be done by having as first step, the making of a copy of

the original, or by having a double copy to begin with. One copy will be "active" and will be read and acted upon; the other will be in a passive state and will merely be copied and recopied at the growing-tip end of the linear array, until the moment for replication occurs.

The precise way in which this is done can vary somewhat. Given a single copy of an initial string, the first step could be to go into a copy routine or replicate mode where the first cell sends a "replicate me" signal to the growing tip final cell, and then transfers (replication) control to the next cell which in turn sends a "replicate me" signal to the newly created final cell, etc. When the last cell of the original copy has sent its "replicate me" signal, control is then transferred back to the left most (non-deleted) cell which now proceeds to go into its regular construction-computation program. This means that our system will require two growth modes: one for replicating itself, and one for carrying out the "normal" constructional-computational activities. It is a matter of convenience to separate or not, these modes.

When the system, in its "normal" mode, has read through and acted upon the original string and encounters the replication of the initial string, "normal" action will not read and act upon this passive genetic copy, but will instead re-create one-by-one at the growing tip, the passive genetic cells. If later the "order" for system replication is given, control is shifted to the first cell of this passive copy, the cell being put in "replicate"mode, and the system proceeds to act as described earlier.

[We have described the switching from "normal" to "replicate" mode as if two-way transmission of information is required. (After the passive genetic material is replicated, we return leftwards to the beginning of the original genetic material, this time treating the genetic material

as being in "normal" mode.) This "leftward signalliñg" can be dispensed

with employing the following technique: during the replicate process we

permit a dual control of activity: two places are being read at the same

time. At the same time that the replicate mode control is reading through

the genetic material, retranscribing it at the growing tip end, the normal

mode control is following behind; no special signal need be sent in a left-

ward direction (except perhaps across one intervening set of cell walls).

The timing and signalling can be controlled by local conditions. Notice

that in circumventing the requirement that a "start" or transfer control

signal be sent "backward" down the full length of the genetic pattern of

cells, we have in turn violated another "ground-rule" of the underlying

monogenic normal system: we now have multiple simultaneous control points:

the string is being read in two different ways, at two different places,

at the same time. The proper cell growth can be made to occur at the tip

end however. The system can be set to operate so that the normal and the

replicate mode activities do not encroach on each other, nor their signals

corrupt each other.

One point might be given further thought: if the genetic material is

Note also that we could have a system in which the genetic material

could be dispersed throughout the cell system (instead of always remaining

contiguous save in the act of retranscribing or replicating). Individual

genetic cells (i.e., in a passive mode) could, when encountered by the

active control process, be recreated singly and separately at the growing

tip end. Depending upon the general mode, normal or replicate, these

cells would be merely retranscribed, or would be both retranscribed

(replicated) and then acted upon (the original passive genetic material

becoming active and taking part in the normal cell-system growth).

One point might be given further thought: if the genetic material is

thus dispersed, will some additional punctuation be required to tell which is the first, and which the final member of the genetic string, or, can a system be set up in which this punctuation is not required?]

The question of the precise conditions which shall result in the replicate routine being started, is very interesting, and should be given some careful thought. We have seen earlier that the general problem of self-initiation of replication in a system of this complexity is not solvable, i.e., if for instance replication is to take place if and only if the computation activity halts, then in general we will not be able to decide for all systems whether replication will take place.

Should replication take place after some fixed finite number of time steps or growth phases? When some other internal criterion is met? When some external, environmental condition is met?

One approach to this problem of the appropriate criterion for replication is to consider what processes the whole system may be carrying out, and ask when it would be useful for such a system to "recreate" itself.

Suppose the system is performing a computation. "Aging" or other "wear-and-tear" on the cell machinery can be thought of as ultimately leading to malfunction. Then we might say that replication is useful because it repeatedly renews the cellular machinery before structural flaws can cause operational-functional errors. (Of course, in the particular linear cell system we have described so far, constant renewal already takes place, but this renewal need not be a permanent feature of the system.) If a computation is to be <u>continued</u> (and not just repeated) by a machine then this renewal of the basic machinery makes sense; renewal, vis-a-vis computation makes far less sense if the new machine merely <u>starts</u> <u>all</u> <u>over</u> <u>again</u>. Except for some value in recomputation for checking purposes, it seems to

make little sense to repeatedly start a computation over again, identically, from the beginning.

[This idea that the offspring machines might carry forward the computations initiated by a parent machine, is very intriguing. Abandoning for a moment the cell-string framework we have been employing, it seems not a bad idea to "pretend" that a Turing machine is capable of useful practical computation, and to consider whether the finite-state control head of a Turing machine might not be replaced periodically, as sort of "preventive maintenance". One can also envisage the replacement of those tapes squares which are frequently used.]

The idea of constant identical repetition makes more sense if we conceive of the system behavior as resulting in the production of needed materials or parts, rather than computations. The larger system (of which our linear cell-system might only be a small part) may require many identical sub-units or sub-processes. The "sensibleness" of this is probably not merely the bias of persons familiar with the notion of industrial production lines; the patterned, repetitive, sub-unit building block nature of much biological construction is well known. Note that our linear cell system with its new cells being created at one end, and old ones being deleted (rendered inactive) at the other, provides a natural production scheme: the deleted cells can, as strings, as individual cells, or as sub-cell parts, be detached at the left to take part in activities elsewhere in the system. If the detachment takes place at the point and time of the beginning of a new replicate phase, then we would have separate copies of a complete cell-system cycle.

9. VON NEUMANN MACHINES AS CELLS AND THE

LINEAR SYSTEM AS ORGANISM

We have presented informally, and then discussed at length, two
cell automaton systems which can be employed to model biological phenomena.
These two systems, the von Neumann two-dimensional cellular system, and
a linear cell system, will continue to be employed in this paper and
will be exploited for their potential applications to biological theory.

Both of these systems posses, in a certain technical sense, universal
computation and construction properties. Thus in a sense the systems are
equivalent and can be applied to the describing of any mechanistic process
of biology.

[We have earlier suggested their use in describing biological phenomena
at the molecular, the cellular, and the multi-cellular level.]

The two systems do have different spatial descriptive properties, and
thus it is reasonable to expect that they will not be equally suited for
describing particular biological phenomena. Thus although there are precise
mathematical senses in which these two systems are equivalent, we may wish
to assign different "natural" interpretations to them. The interpretations
upon which we are going to concentrate are by no means the "only" or even
the "only natural" ones. I believe however that the interpretation will be
fruitful and revealing.

The first association is between machines of the von Neumann cellular
system and individual biological cells and their intra-cellular activity.
That is, we wish to view a von Neumann cellular machine as a single bio-
logical cell, and whenever this is the interpretation principally in mind,
will call it a von Neumann Cell. Self-duplication of a von Neumann machine
thus becomes biological cell duplication, the instructional tape becomes cell
DNA, etc.

For a second association, we wish to concentrate on the interpretation of a linear cell system as a multi-cell organism. Problems of the creation, growth, differentiation of particular cell systems and their individual cells, become problems of the growth, organization, and development of multi-cell organisms.

The two automaton systems which we have heretofore described "in parallel" are thus now to be seen as hierarchical. We wish to consider formally the morphogenesis of organisms whose individual cells are formally clusters of von Neumann system cell-components.

In order for the von Neumann Cells ("machines" in the earlier sense) to serve as the cells of the linear cell system, they must possess not only self-duplication properties but some additional properties determining the precise order and kind of offspring produced, and provision for changes in growth and activity repertoires as a consequence of the number and nature of neighboring Cells. Included in the repertoires we will want to allow for branching growth of Cells, and for transmission of ("chemical" or "electrical") impulses between Cells.

The von Neumann cellular system will be employed then as the vehicle for explicating single cell and intra-cell problems, eg. cell duplication, DNA replication, virus invasion, genetic mutation, etc., while the linear cell system will be employed as the vehicle for explication of problems of morphogenesis.

In the next subsection we discuss problems and directions of research for systems of von Neumann Cells. In forthcoming sections we will further discuss linear cell organisms.

9.1. Von Neumann System as a Duplicating Cell. If we view a von Neumann cellular machine as a mathematical expression of the activities of a single

biological cell, the problem of the form and use of "genetic" construction tapes, (the "DNA" of the Cell) is central.

Q1: What sorts of Cell duplication are possible? What sorts of interaction between DNA code and Cell form are possible?

Let us begin by describing an "ideal" form of Cell duplication in the von Neumann system. A von Neumann Cell can be designed which is capable of universal construction. The Cell is equipped with a tape (DNA) the contents of which consists of two copies of the instructions for the construction of a von Neumann Cell. The instructions include copying the instructions into the new Cell (by using the second of the two original copies to make two copies to be inserted in the daughter Cell). The daughter Cell can be (and for this first case we shall assume is) an exact duplicate of the original Cell, including tape contents. The off-spring Cell is constructed to remain passive until a stimulus pulse is sent to it by the parent Cell (this is the final instruction in the genetic tape--with the possible exception of an instruction to sever connection with the daughter cell) after which the child Cell begins to operate autonomously. This Cell self-duplication we shall call the von Neumann standard form of Cell self-duplication. This von Neumann standard self-duplication is usually conceived of as being carried out in a clear region of the operating space by a "constructing arm". This constructing arm is (at least for cell-duplication) very unbiological. See Figure 9.1.

Q2: Is there a fruitful revealing form of von Neumann Cell duplication achieved by means closer to the naturally biological: a modelling of cell fission, perhaps?
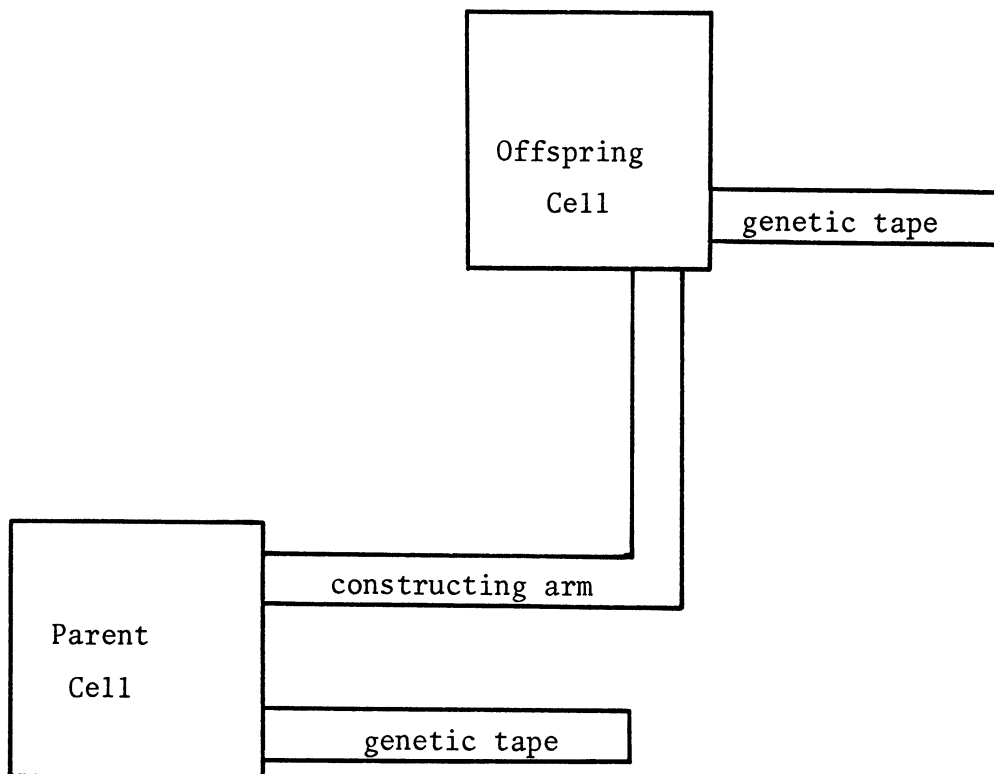
Schema of Von Neumann "Standard" Self-Duplication

Fig. 9.1.

Of the many questions that a mathematician may ask of classes of objects (objects such as machines, strings, theorems or putative theorems, etc.) the mathematical formalist has two favorite ones: can one algorithmically tell whether an object is a member of the class, and can the members of the class be listed algorithmically. That is, is the class recursive? Is the class recursively enumerable?

We have already several times noted that in general we cannot tell whether a von Neumann machine (a Cell, in the terminology of this section) will duplicate. That is, the class of von Neumann self-duplicating Cells is not recursive.

Q3: Is the class of von Neumann Cells recursively enumerable? Can we list the self-duplicating Cells?

This problem of listing is simpler than the problem of deciding since if we can decide, we can list, but if we can list we can not necessarily decide.

If we limit our notion of self-duplication to the von Neumann standard form we described briefly above, then we can indeed enumerate the self-duplicating Cells. We could do this by the following technique: 1) enumerate all the Cell types (we can do this by enumerating first all the one cell-component Cells, then all the two cell-component Cells, etc.), 2) set each Cell in action, 3) after each action, check to see whether the newly produced Cell (if any) is identical to the original Cell (this checking is a finite process), 4) if any offspring Cell proves to be identical to the parent Cell, (including genetic tape information) place the parent Cell in the self-duplicating class.

By this technique we can enumerate all von Neumann standard form self-duplicating Cells.

[Notice that not only are there Cells which are not self-duplicating, but there are Cells which can not be constructed even by designing special constructing Cells for the job. This fact was pointed out by Moore and by Myhill where it is shown that there are "Garden-of-Eden" configurations possible in the von Neumann cellular system; a Garden-of-Eden configuration of cell-components is one that can exist only at an initial moment of time, active cell-components within it immediately changing it. A Cell trying to produce some such configurations could not withdraw itself from the construction site fast enough to avoid "contaminating" (changing the behavior of) its own creation. For this reason, the standard form of self-duplication produces a _passive_ offspring which only at the last moment is stimulated into life.]

Thus, even if the parent Cell wanders around over the landscape doodling cell components here and there, _if_ eventually it _is_ a standard self-duplicating Cell, we will be able to recognize this when it occurs.

[The effect, the usefulness of standard self-duplication could of course be swiftly vitiated if the parent and offspring cell immediately begin to conflict. Self-duplication would have, however momentarily, taken place.

In explicating this "standard" form of Cell self-duplication in the von Neumann Cellular System, several conventions have consciously been employed, principally as an aid to explication. The parent Cell is generally thought of as a _contiguous_ block of stimulable component cells. There is a single, readily recognizable tape unit attached at the lower right corner of the parent Cell. There is a single, readily recognizable con-

structing arm extending from the upper right corner of the parent Cell.
The Cell operates serially rather than in parallel. The offspring Cell is
pictured as being constructed in a specific location above and to the right
of the parent Cell. The offspring Cell is compact and contiguous (like
the parent). The point at which the constructing arm of the parent leaves
off and the offspring Cell begins is clear (or only momentarily ambiguous
during some construction phases).]

This form of duplication is, of course, not a required form; it was
chosen as an expository device. Below some discussion of "nonstandard"
forms of duplication is given.

9.1.1. Nonstandard Duplications. Let us now begin the discussion of
cases where although we should like to be able to tell whether we have a
self-duplicating Cell, or whether self-duplication has taken place we
may not be able to do so.

The "standard" self-duplication of Cells in the von Neumann cellular
system, (which we described in the last section) is only one way in which
the duplication process can take place. The standard form of duplication
disguises many difficulties which can arise in the general case of dupli-
cation in the von Neumann system. The necessary activities can take place
in many ways. For that reason, it may be useful to broaden one's view
considerably, and recall some of the ideas presented in Section 2 where
structures possessing active sites were discussed as a very general model
of the biological processes of interest.

From this point of view, a Cell in a von Neumann cellular system
consists of some structure of component cell elements, and some number of
active sites. At these active sites, elements can be created, can be strung

together, can cause internal changes to take place. The elements of structure produced at one site can effect the machine by impinging on other active sites; they can combine to produce larger autonomously acting structures which in turn can attach themselves to the original structure, or produce elements which will play a structural or informational role vis-a-vis the original machine.

Thus, in a von Neumann cellular machine: there can be multiple sites of activity; multiple simultaneous sources of "genetic information". The Cell need not necessarily be compact, or even contiguous. The "genetic reading" operations, and the "offspring constructing" operations could be thoroughly intermingled. The parent and the offspring Cells could be intertwined so thoroughly that particular activities could not easily be distinguishable as parental or offspring. The offspring Cell need not be contiguous. The offspring need not be produced in any "canonical" location. The parent Cell might wander in desultory fashion, constructing all sorts of curlicues before beginning offspring construction. The offspring might be created in an envelope completely surrounding the parent. The offspring might incorporate smaller or larger parts of the parent into itself. The parent might separate itself from its completed offspring at one moment, and begin destroying it the next; on the other hand bringing offspring to maturity can also have grievous effect on parents.

The offspring Cell might, with some degree of autonomy, begin operation before it had been completely constructed. It might carry out some of its own construction. Both the parent and the offspring Cell might send materials or signals into themselves or their environment, which could return to affect them further.

And all these things and more could be going on at the same time.

Q4: What is the <u>least</u> we would accept as an instance of self-duplication? Can we capture precisely the notion of self-duplication? Can we produce a formal definition of self-duplication that would gain general acceptance?

In any case the possibilities are so numerous, and a formal definition so elusive, we will at this point confine ourselves to consideration of cases quite close to those of the standard format. Even so, there are many difficulties.

In von Neumann's scheme for machine self-duplication, a parent machine (or Cell) has two tapes (or two sections of the single tape) available to it, an "active genetic" tape which consists of instruction to be acted upon to produce a "child" machine, and a "passive genetic" tape, the contents of which are to be copied (twice) into the child Cell. We wish to contemplate the range of possible characteristics of four elements of our cell biological system model: the <u>Parent Cell</u> proper, the <u>Active Genetic Tape</u>, the <u>Passive Genetic Tape</u>, and the <u>Offspring Cell</u> (or Cells). For convenience let us abbreviate these elements PC, AGT, PGT, and OC, respectively.

Let us list a few of the things which can occur.

1) The PC is given an AGT. PC and AGT are "suitable to each other" (PC is capable of carrying out tape instructions, and this includes the AGT). AGT has the instructions for an OC; OC is produced.

This is an instance of a Cell producing a Cell.

2) If PC is also given a suitable PGT and the PGT information is transmitted to OC, <u>and</u> OC and the PGT are suited, then further offspring, OOC (Offspring of Offspring Cell) may be produced.

3) If PC is given an AGT where AGT is instruction for building a

OC <u>and</u> where PC and AGT are suited and where PC and OC are identical, the

PC duplicates itself (minus of course the AGT, which may have been

expended or altered in the course of OC's construction).

4) "Complete" duplication takes place if PC is given an AGT and a

PGT where AGT acted upon yields an OC which is identical to PC and PGT

is a copy of AGT. PC copies PGT twice into OC. OC is now completely

identical to what PC was. "Complete" duplication has by this means taken

place.

[If PC and OC must be <u>distinct</u> and <u>identical</u> for self-duplication to

have taken place, then we can check to see if self-duplication has indeed

taken place. The class of self-duplicating PC, under such stringent

definition, is thus recursively enumerable.]

J. von Neumann showed that there is a PC and there are suitable genetic

tapes, such that the PC is universal, that is that it can proceed to con-

struct anything that <u>is</u> constructible in the system. Among the things <u>not</u>

constructible in a string sense are Garden-of-Eden configurations etc.;

among the things not constructible in a weak sense are certain Cells with

particular properties of timing or form. This latter is not as great a

limitation as might at first have been supposed. (Holland and Mukhopadyay

have considered what behaviors and what computations are possible.)

There are PC which when given a suitable AGT will produce an OC which

is identical to PC, and <u>where the PC is not a universal constructor.</u> Some

of these less-than-universal constructing systems are quite trivial ("domino

falling against domino").

Q5: How much can be accomplished by less-than-universal constructors?

5)   A PC can be given an AGT and PGT where AGT is the instruction
to produce offspring composed "principally" of AGT and PGT pairs which
are in turn suitable, to be acted upon by other PC.   This is a scheme
of "virus" invasion.

In 5, the offspring do not contain complete construction capability
but must employ the complete construction machinery of a suitable PC.

It is also possible that a PC possesses complete construction machinery
and possesses tapes suitable to it, but does not act upon these tape
instructions; PC is inactive, or lacks some stimulus signal.

Q6:   Is the set of PC which would be self-duplicating if stimulated
recursively enumerable?

Q7:   Is the set of PC which if given suitable AGT and PGT would be
self-duplicating, recursively enumerable?

We have heretofore largely contemplated PC which have been "stripped
down" to their construction machinery.   PC may of course possess additional
properties not directly relevant to their machinery for constructing off-
spring.   These additional properties can in particular govern the number,
time, and place of offspring construction; the relation to the environment
and to other Cells in the vicinity, including sending and receiving of
electro-chemical messages and the production or absorption of materials
(eg. the production of signals or components which other Cells may employ
or act upon).   In particular, Cells will have a repertoire of behavior
which will enable them to take part in morphogenetic activities at the
multi-cell level.

These properties auxiliary to the duplication process can be extremely complex. If we assume that an organism starts from a single cell and that the genetic information remains stable, then these auxiliary properties become the means by which cells are differentiated and distinguished. Thus in many cases of interest PC and OC will be genetically identical but not at all identical cell-component-wise.

6) It is not necessary, of course, that AGT and PGT presented to a PC be identical. PC acting upon AGT may produce an OC not identical to itself (not identical in any of a large number of respects). OC is given two copies of PGT. If OC and PGT are suitable to each other, OC can act upon PGT and can, for example, produce a copy of itself, or even of PC. In the latter case it is reasonable to say that PC is self-duplicating, but with a generation lag. From this point on the Cells should "breed true".

Q8: Can alternation of form by generation be perpetuated?

7) The anomaly described in 6, can be made to persist one more generation. The PGT given PC may be the description not of PC or of OC but of OOC. Only from OOC on does the strain breed true. Should PC be termed self-duplicating?

Q9: What is maximum number of generations over which such anomalies can be made to persist? What is the maximum n such that $O^n C$ is identical to PC but $O^{n-1}C$ is not identical to OC (cf. "busy beaver" problem of Rado)?

We thus have cases in which Cells are to be distinguished by their

genetic material and cases where they are to be distinguished by their auxiliary properties. We list some cases below:

a) PC and OC are spatially distinct and are completely identical in form.

b) PC and OC are spatially distinct, identical in external behavior (eg. step-by-step growth actions), different in form.

c) PC and OC are spatially distinct, identical in "computation-construction" (ultimate results the same but methods of producing the results may radically differ), different in form.

d) PC and OC not spatially distinct; PC and OC perform the same construction-computations, but one cannot always tell where PC leaves off and OC begins.

e) PC and OC spatially distinct; they behave identically in respect to genetic-construction tapes, but their auxiliary properties differ.

In order to explore problems such as the above, it may be useful to employ computing machinery. A computer simulation system allowing the expression of von Neumann Cells, and permitting easy and swift change in Cell properties would undoubtedly be of great assistance in suggesting conjectures, and in evaluating those conjectures.

[This paper will continue with a discussion of Linear System Organisms, Problems of Linear System Organisms, and Production Systems for Linear Organisms employing ideas drawn from theory of formal grammars.]

Polanyi, M., "Life's Irreducible Structure", Science, 160, 3844, June 21, 1968, 1308-1312.

Post, E. L., "A Variant of a Recursively Unsolvable Problem", Bulletin of American Mathematical Society, 52, 264-268, 1946.

Rado, T., "On Non-Computable Functions", Bell System Technical Journal, 41, 1962.

Scott, D., "Some Definitional Suggestions for Automata Theory", Journal of Computer and System Sciences, 1, 187-212, 1967.

Shannon, C., "A Universal Turing Machine with Two Internal States", Automata Studies, Princeton, 1956.

Shepherdson, J. C. and H. E. Sturgis, "Computability of Recursive Functions", Journal of ACM, 10, 217-255, 1963.

Stahl, Walter R., "Algorithmically Unsolvable Problems for a Cell Automaton", Journal of Theoretical Biology, 8, 371-394, 1965.

Smith, H. R., "Simple Computation-Universal Cellular Spaces and Self-Reproduction", Record of Ninth Annual Symposium on Switching and Automata, 269-277, 1968.

Thatcher, J., "Universality in the von Neumann Cellular Model", The University of Michigan Technical Report, 1965. To appear in Essays on Cellular Automata (ed. A. W. Burks), University of Illinois Press.

Thompson, D'A. W., Growth and Form, Macmillan, 1942.

Ulam, S., "On Some Mathematical Problems Connected with Patterns of Growth of Figures", Mathematical Problems in the Biological Sciences, AMS 215-224, 1962.

von Neumann, J., Theory of Self-Reproducing Automata (ed. A. W. Burks), University of Illinois Press, 1966.

Waksman, A., "A Model of Replication", JACM, 16, 178-188, 1969.

Wang, Hao, "Dominoes and the AEA Case of the Decision Problem", Mathematical Theory of Automata, Polytechnic Press of Polytechnic Institute of Brooklyn, 23-55, 1963.

Wang, Hao, "Remarks on Machines, Sets, and the Decision Problem", Formal Systems and Recursive Functions (ed. Crossley and Dummett), North-Holland Publishing, 304-320, 1965.

Wang, Hao, "Tag Systems and Lag Systems", Math. Annalen, 65-74, 1963.

Whitehead, A. N., An Introduction to Mathematics, Oxford University Press.

BIBLIOGRAPHY

Arbib, M. A., "Automata Theory and Development", Journal of Theoretical Biology, 14, 131-156, 1967.

Arbib, M. A., "Simple Self-Reproducing Universal Automata", Information and Control, 9, 177-189, 1966.

Apter, M. J., Cybernetics and Development, Pergamon Press, 1966.

Barricelli, N. A., "Numerical Testing of Evolution Theories", Acta Biotheoretica, 16, 69-98, 1962.

Berger, R., "The Undecidability of the Domino Problem", Harvard Computation Laboratory Report, 1964.

Bonner, J. T., Morphogenesis, Princeton University, Princeton, N. J., 1952.

Cocke, J. and M. Minsky, "Universality of Tag Systems with P = 2", Journal of ACM, 11, 1, 15-20, 1964.

Holland, J. H., "Hierarchical Descriptions, Universal Spaces and Adaptive Systems", University of Michigan Report, August 1968.

Lieblein, Edward, "A Theory of Patterns in Two-Dimensional Tessellation Space", Ph.D. Thesis, University of Pennsylvania, 1968.

Lindenmayer, A., "Mathematical Models for Cellular Interactions in Development", Journal of Theoretical Biology, 18, 280-299, 300-315, 1968.

Mc Naughton, R., "On Nets Made up of Badly Timed Elements, Part I: Slow but Perfectly Timed Elements", Moore School, University of Pennsylvania, 1961.

Minsky, M. L., "Recursive Unsolvability of Post's Problem of "Tag" and Other Topics in Theory of Turing Machines", Annals of Math, 74, 437-454, 1061.

Minsky, M. L., "Size and Structure of Universal Turing Machines Using Tag Systems", Recursive Function Theory, Symposia in Pure Mathematics 6, AMS, 1962.

Moore, E. F., "Machine Models of Self-Reproduction", Mathematical Problems in the Biological Sciences, AMS, 17-33, 1962.

Mukhopadhyay, A., "Representation of Events in the von Neumann Cellular Model". Journal of ACM, 15, 4, 693-705, 1968.

Myhill, J., "The Converse of Moore's Garden of Eden Theorem", Proc. AMS, 14, 685-686.

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
LOGIC OF COMPUTERS GROUP The University of Michigan Ann Arbor, Michigan	Unclassified
	2b. GROUP

3. REPORT TITLE

FORMALISMS FOR LIVING SYSTEMS (Part I)

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report

5. AUTHOR(S) *(Last name, first name, initial)*

Laing, Richard A.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1969	92	33

8a. CONTRACT OR GRANT NO. DA-31-124-ARO-D-483	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	08226-8-T
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	U. S. Army Research Office (Durham) Durham, North Carolina

13. ABSTRACT

Various formalisms for expressing the structure and behavior of biological systems are proposed and discussed. The formalisms include tiling or "domino" systems, the von Neumann cellular automaton system, and several systems based on the monogenic normal systems Tag and Lag. There is some discussion of desirable characteristics of models for biological theory.

DD FORM 1473 1 JAN 64

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
automata						
cell machines						
von Neumann						
biological theory						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through

_____ ."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through

_____ ."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through

_____ ."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.