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

Technical Report

ARTIFICIAL ORGANISMS
AND AUTONOMOUS-CELL RULES

Richard Laing

with assistance from:

ARTIFICIAL ORGANISMS AND AUTONOMOUS-CELL RULES

Abstract

Developmental rules in which a cell may differentiate or propagate as a consequence of its internal state only are shown to be sufficient to produce any given organism cell-structure. Rules in which a cell may differentiate or propagate as a consequence of its internal state *and* the number of other cells impinging on it, are shown to be sufficient to produce an organism possessing universal computational behavior.

ARTIFICIAL ORGANISMS AND AUTONOMOUS-CELL RULES

1. Introduction

Although it is certain that the cells of developing biological organisms are in fact capable of considerable amounts of intercommunication, it is not at all clear how much intercommunication is actually employed in producing organism form. It is therefore an interesting theoretical problem to establish first how much cell intercommunication is *necessary* in order to achieve certain organism forms.

The structures of simple organisms can be represented by finite graphs of points and lines where the points stand for cells and the lines stand for contiguity relationships between cells. The points and lines of such graphs might be labelled to represent respectively the internal states of cells and the presence of specific kinds of contiguity relationships to other cells. In a formal approach to biological development we can then consider how, starting from a single initial cell, complex cell-contiguity graphs reflecting organism structure, might be produced by systems of simple rules of growth. Autonomous-cell re-write systems (in part akin to the context-free grammars of the formal linguist) have sometimes been suggested for this purpose.

In autonomous-cell rules (as we shall employ them here) the condition of a given cell may be re-written as a consequence of local characteristics of the cell (such as the internal state alone of the cell, or the internal state along with certain local contiguity properties) and *not* (for example) as a direct consequence of the internal states of neighboring cells. Thus, autonomous-cell rules reflect a minimum of intercellular communication: a cell may act upon the knowledge that it has some number of

neighbor impinging on it but it does not know directly anything of the internal condition of cell-neighbors.

One result of applying an autonomous-cell re-write rule is merely to change the state of a given cell. We will call such a rule a *differentiation rule*. Another result of a rule application can be to produce one or more new cells (in specified states); we will call such a rule a *propagation rule*. In addition to specifying the states of existing and newly created cells, differentiation and propagation rules must also implicitly or explicitly specify what connections to other cells are to be retained or lost. These connections to be revised are restricted to those already locally present at the cell to which the rule was applied. Thus, a (binary fission) autonomous-cell propagation rule applied to a cell converts it to two successor cells; the possible connections which these two new cells may make to other cells are restricted to a connection to each other and connections, by each of the new cells, to any or all of the cells the parent cell was connected to. In the case of a differentiation rule (as we define it here) some of the existing connections may be lost, but there is no creation of new connections. Notice that the rules do not permit the dissolution of a cell, although the rules will permit the complete detachment of a cell.

We now define some even more limited autonomous-cell rules.

A *restricted autonomous propagated rule* will produce any number of new cells with prescribed connections, only *one* of which new cells is in a state which can result in the propagation of further new cells. A *restricted autonomous differentiation rule* will never place a non-propagating cell in a propagating state. An autonomous cell system of rules in which both the propagation and the differentiation rules are of the restricted sort we will call a *restricted autonomous system.*

The autonomous and restricted autonomous systems we have outlined above may seem to possess quite modest powers. It is therefore interesting to see whether such rules are sufficient to account for the forms and behaviors assumed by biological organisms.

2. Construction of Arbitrary Cell Structures

Autonomous-cell rules can be employed, starting from a single initial cell, to construct, in deterministic fashion, any single specific (finite) organism cell-contiguity structure whatsoever.

One way in which this may be achieved is the following: to construct any given finite cell-contiguity structure of n points we first construct a *uniquely* labelled completely connected cell contiguity structure of n points. This is done by applying a set of unique propagation rules, beginning with the given initial cell, n-1 times. At each application of a propagation rule a uniquely labelled cell (to which the propagation rule applies) is re-written as two, uniquely-labelled cells. These daughter cells are, at each rule application connected to each other and to each cell their parent cell was connected to. For notational convenience only, unique labels are also given to the contiguities as they are created.

From this uniquely labelled complete contiguity structure of n cells we now wish to create the final desired labelled structure of n cells. First (in any manner) determine an association between each unique cell label and the corresponding cell and label of the desired final structure. Then employ unique differentiation rules (applicable one each to each of the uniquely labelled cells) that will re-write the unique cell label to the desired final cell label and at the same time specify the selective deletion of any superfluous connections to other cells. The result will be the

desired cell contiguity structure.

As an "explanation" of development such a scheme leaves a great deal to be desired. The notion of creating a completely connected proto-organism cell-contiguity graph, and then severing the non-organism contiguities hardly reflects the course of most real biological development and there are of course other objections which might be made to this scheme, as an explanation of development. (See the critique in Laing and Arbib (1972)).

Is there a way of employing autonomous-cell rules so as to produce a desired n-cell organism contiguity structure without the course of development passing through the n-cell completely connected stage?

Theorem: For every given finite cell-contiguity structure of n cells, there is a system of autonomous cell rules which will deterministically yield precisely the given structure. Moreover, this can be done

1) starting from a single initial cell,

2) using *restricted binary fission* propagation rules only,

3) without (for a desired structure of n cells) during development producing a *completely* connected structure of n cells (which is not also the desired final structure),

4) employing only "pure" differentiation rules (in a pure differentiation rule only state change can occur; there are no deletions of cell contiguity relationships); indeed separate differentiation rules may be dispensed with entirely.

Note that 1) taken with 2) means that there need never be more than a single propagationally active cell in the developing organism structure graph.

Proof: Given any finite cell structure:

1) *Uniquely* label each of the cells and (again for notational convenience) the connections and introduce into the autonomous re-write system

one differentiation re-write rule for each of the unique cell labels such that where A is an original cell-state and K is its new *unique* label, there is a differentiation rule K → A. (We shall later see agreeing with condition 4) of the theorem, that this set of differentiation rules is eliminable.)

2) The given cell structure may consist of isolated cells and one or more connected components. Choose any isolated cell and merge it with any other isolated cell. Uniquely label the new combined cell and introduce into the autonomous system of rules a propagation rule which has as its antecedent the new uniquely labelled cell and as its consequent the two original, properly labelled cells. Since the cells were originally isolated, the newly introduced propagation rule will not preserve the connectivity between the two consequent cells. Continue in this fashion until there is only one isolated cell. If there are no further organism components go to step 6).

3) If there are further connected organism components, merge the remaining isolated cell with a cell of any connected organism component. Uniquely label the newly formed merged cell and introduce a propagation rule which completely severs the connections to the cell which will be isolated, and retains the appropriate connections to the cell which is in the connected component.

4) Take the newly merged cell (call it D) of the connected component and merge it with any other cell (call it E) of the component to which it is *directly* connected. Uniquely label this new merged cell (call it C). In the cell structure, connect all of D's connections to C. Introduce a propagation rule in which

the rule will also specify the connections D and E and which

the rule will also specify the connections D and E had to other

points before the latest merging into C.

5) Continue in this fashion in the connected component and (as in 3))
with any remaining connected components. Eventually a single
cell will remain.

6) Introduce the final single cell itself into the autonomous cell
system as the desired initial fertilized cell.

This completes the main parts of the proof. In the procedure outlined
above we have employed only pure differentiation rules along with propagation
rules in which only one of the two offspring points or cells is free further
to propagate; that is, all of our propagation rules have been of the binary
restricted sort. Also, since the system we have produced begins with a
single initial cell capable of propagating, and our *restricted* rules preserve
this property, we have employed only a single "live" propagating cell in the
developing structure. Notice also that development proceeded by a more
"natural unfolding" than by the "completely connected contiguity structure
followed by deletion" technique outlined earlier.

In condition 4) of the theorem we noted that we can eliminate diff-
erentiation rules entirely. This is accomplished by having the appropriate
cell states assigned step-by-step in the course of applying the propagation
rules. The only problem that arises in this is a notational one. By
assigning unique labels to the cells and retaining these labels until all
propagation is completed we can in our proof, use them to name the contiguities
that are to result from rule application. If we should, in the course of
propagating new cells, place cells in their final differentiated state
however, we may lose this notational convenience. For it is quite possible

that many cells may ultimately come to be in the same final differentiated state. Thus using these states to name contiguities may produce ambiguity.

Note however that rule application and contiguity assignment does not depend on knowing which cells are adjacent to the cell under consideration. An autonomous rule merely assigns contiguities from those connections *locally present;* the rule is indifferent to the ultimate destination of the connection.

We now given an example of obtaining a set of autonomous rules for constructing a given cell contiguity structure. In this example we employ the notational convenience of specifying connections (labelling lines) by giving the labels of the contiguous cells. Thus, 6' → 6(5'), 5'(6) means that cell 6' is to be re-written as two cells, 6 and 5' and that they are to be connected to each other (6 is connected to 5' (and 5' is connected to 6)).
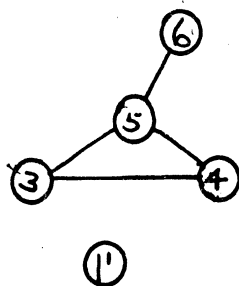
Example

1. Given the cell-contiguity structure:
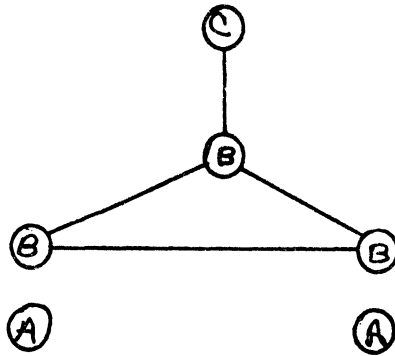
2. We uniquely label the cells and set down the (pure differentiation) rules which will enable us to recover the original structure.
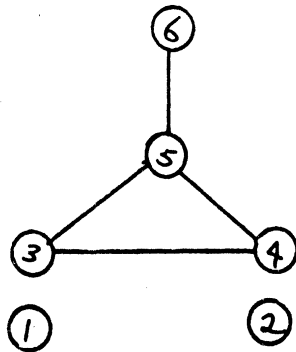
$$1 \rightarrow A$$
$$2 \rightarrow A$$
$$3 \rightarrow B$$
$$4 \rightarrow B$$
$$5 \rightarrow B$$
$$6 \rightarrow C$$

3. We merge the two isolated cells and set down the rule which will enable us to recover them (and thus restore the structure of 2.).
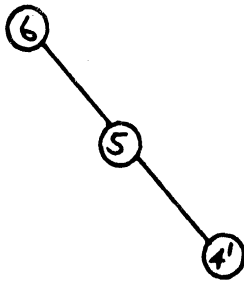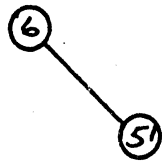

$$1' \rightarrow 1,2$$

4. Continuing:
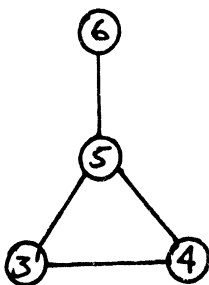
$$3' \rightarrow 3(5,4),1'$$

5)

$$4' \rightarrow 4(5,3'), 3'(5,4)$$

6)

$$5' \rightarrow 5(6,4'), 4'(5)$$

7)

$$6' \rightarrow 6(5'), 5'(6)$$

The single initial cell   of our autonomous cell system is to be given the label 6'.

3. Universal Artificial Organism Behavior

We have shown (in Section 2) that any single particular artificial organism cell-structure (finite point-labelled graph) can be constructed using only simple (restricted autonomous-cell) re-write rules. In these organisms only one cell at a time was capable of propagating new cells, and indeed in this formulation no cell once having ceased to be a propagating cell ever regains this property. We need not however have restricted ourselves to such special organism systems. Indeed it is biologically more interesting (though also perhaps mathematically less challenging) to consider organisms which can possess numerous active cites (or even multiple *simultaneously* active sites). For a living organism (even an artificial one) need not consist of a mere static structure of cells. We should therefore like to consider not only what *structures* but what *behaviors* are possible for artificial organisms growing and changing according to simple rules.

What we will show in this section is that employing only autonomous-cell rules of differentiation and propagation we can produce an artificial organism in whose dynamic behavior any effective computational process can be represented.

Theorem: Employing autonomous-cell rules only, an artificial organism can be constructed whose growth and change simulates the computation of any Turing machine.

A Turing machine (Turing (1957)) is an abstract computing device consisting of a finite-state "reading-head" mounted on an indefinitely extendible tape marked into squares. A Turing machine operates by examining the symbol on the tape square under the reading-head and then, as a consequence of the state of the reading-head and the particular tape symbol being read, may change the tape symbol under scan, change internal state,

move one square right or left. A Turing

machine computes functions by acting upon the argument of a function inscribed

initially on the tape, and printing out the value of the function on the tape.

It is generally accepted that a Turing machine can carry out any effective

procedure: that is, that any calculation that intuitively can be carried

out in rote fashion is computable by Turing machine.

It is known that "universal" Turing machines exist: Turing machines

which when presented on their tapes with both an argument of a function

and the *description* of a Turing machine which computes the function, will

proceed to carry out the computational task for the given argument by

following the instructions implicit in the machine description given on the

tape. Such universal "simulating" machines need not be terribly complex;

universal computational behavior is possible for Turing machines having

fewer than 10 reading head states and 10 different tape symbols. (The

present minimum state-symbol product is 28: 4 symbols, 7 states (Minsky (1961)).

A Turing machine may be said to distill much of the essence of information

processing, whether in a natural or artificial system. In a rudimentary

but (adequate) fashion a Turing machine receives stimuli from its environment

(tape),   selects, records, and transforms the stimuli (at the same time

transforming itself by these actions) and in turn acts upon its

environment.

It has long (at least since the mid-1950's) been common knowledge among

automaton theorists that the action of any Turing machine (including thus

universal Turing machines) could be carried out by an indefinitely extendible

string of finite automata. Each of the automata would embody first of all

the internal finite state structure of the Turing machine reading-head;

each automaton would also possess certain additional storage capabilities

which would allow it to record which tape symbol it would, for its position

in the row of automata, have under scan; also, certain additional capabilities

must be present to allow each automaton to recognize whether it is the

automaton presently playing the role of active Turing machine reading-head

(plus symbol under scan), or whether it is playing the role of passive

reading-head (and tape symbol). Finally each automaton must possess the

inter-communicating capability to cause neighbors to the right or left

to change their states and to assume the role of active reading-head, while

at the same time relinquishing the active role itself.

Artificial organism systems composed of cells with the same capabilities

as the automata described above can, in the same fashion, carry out universal

computation. In this paper, however we have been restricting our attention

to a very simple sort of organism system in which differentiation and propagation

activity occurs only as a consequence of very local conditions. In these

autonomous-cell systems a cell cannot (for example) communicate a change of

state directly to a neighbor. This limitation and other paucity of means

in autonomous-cell systems will sometimes oblige us to employ a *series*

of actions to accomplish what, in more powerful systems, would be a single, simple

and straightforward action. An informal proof of the theorem stated at the

beginning of this section (that artificial organisms constructed exclusively

by autonomous-cell rules, are capable of exhibiting any Turing machine

computation, i.e., universal computational behavior) is now given.

Since by our earlier results any single particular fixed configuration

of cells can be constructed employing autonomous-cell rules alone we can

obtain an initial connected chain of cells, of any desired length, the cells

in any desired state. We will want each cell of this chain to be able to

act like a combined reading-head and (possibly marked) tape square of a

Turing machine. That is, the internal state of each cell can be viewed as

falling into two parts: one part records the present state of the reading-head, and one part stores the state of the tape square which would be under scan at that point. The argument of the function to be computed will be set in the tape state portion of successive cells during the initial construction phase. One of the initially produced cells will be in an "active" state; that is, a re-write rule will be applicable to it. The other cells will be in quiescent states; that is, no differentiation or propagation rule is presently applicable to them, but (owing to later changes in their local conditions) they may yet undergo differentiation and propagation. (Note that we no longer have a *restricted* autonomous-cell system since the rule that a cell once having ceased to be a propagating cell never again undergoes propagation, need not be obeyed here.)

To show how Turing machine behavior (including universal computational behavior) is possible for this artificial organism, we now concentrate on the actions of the single active cell. This active cell represents the Turing machine reading-head actively scanning a tape-square. As in the Turing machine, the combination of reading-head state and tape state (in this organism case both internal to the cell) may produce state change, tape symbol change, and possibly a shift of attention right or left.

If no shift right or left is to be made, then clearly autonomous-cell differentiation rules will suffice to change the cell internal state (that is, change the reading-head part of the total internal state, or the tape square part, or both). Transferring active cell status as well as that part of the internal state representing reading-head status, to a neighbor cell poses the only real problem for representing Turing machine behavior in these limited systems.

Recall that part of what a cell can know and act upon (i.e., what an autonomous-cell/application rule is conditioned by) is the number of cell connections impinging on the cell. We employ this property to communicate the desired information. An active cell in reading-head state k which is to "transmit"

knowledge of this reading-head state to a neighbor cell, (as well as the fact

that the neighbor is also to assume the active role) will have a rule applicable

to it which will cause it to divide into a number k+1 of new cells.

One of these cells is to play the role of old reading-head and it

consequently remains connected to the original right and left hand neighbors

and assumes one of the quiescent states, while retaining a record of the

tape status at its location. The remaining k cells are to be connected to

the neighbor cell and to no other of the possible cells. Whenever a cell,

previously quiescent (having no rule applicable to it) has excess cells appended

to it, there *will* be a rule applicable to it. Application of the rule will

cause the cell to differentiate, assuming a reading-head state k, and at

the same time sloughing off the k new formed cells attached to it. The cell

will now be in computationally active status and will proceed to act upon

the reading-head state and tape-state information it possesses. (It should

be noted that the number of different reading-head states to be transmitted

is fixed and finite, and consequently no rule propagating an indefinitely

large number of cells at one step (or counting off an indefinitely large number

of repetitions of binary propagations) will ever be required.

Since with the above repertoire of actions the behavior of any Turing

machine can be represented, and since some Turing machines are computational

universal, it follows that there are autonomous-cell artificial organisms

also capable of universal computational behavior.

Although the above remarks constitute the essential features of the

desired proof, we will mention briefly some peripheral matters which must

be taken into account.

In order to carry out the computational activity, the string of cells may have to grow indefinitely long, and we have not described the conditions under which the initial string of cells will be extended. This can be handled in several ways, the simplest of which would be to impose continuous growth at the two ends of the string: a cell finding itself with only one neighbor will grow a neighbor to fill the empty location, this activity proceeding concurrently with the computational behavior. The rules can also be revised so that new growth takes place only when necessary to carry out the desired computational beahvior. For example, a cell "knows" it is an end cell because it possesses only one neighbor connection; it therefore can assume a special status so as to record this fact. If it should have control of the computational behavior transferred to it, its altered state will give it the "knowledge" required to take the missing neighbor into account before differentiating so as to assume the proper internal state. If control is to be transferred to the (missing) neighbor, the active cell will first construct the required cell.

That an "organism" is capable of universal computation may not seem very relevant to problems of biology; the naturally occurring behavior of organisms may appear to be not at all like mathematical calculation. It is generally agreed however that the notion of "Turing-computable" characterizes precisely the intuitive notion of "effective procedure". An effective procedure is one which can be carried out in a purely mechanical fashion. Thus, if one believes (as seems reasonable) that it is useful to try to describe biological phenomena in mechanistic terms, then formal systems such as our autonomous-cell systems, provide a precise vehicle for so-doing.

Although we have here showed how a very simple formal biological system can yield any organism form and any organism behavior, it clearly is not

true that all organism development and all organism behavior have thereby been explained. In our simple autonomous-cell formal systems we have merely proffered several (not very convincing) explanations for organism growth and behavior. In the next section we discuss some of the adequacies and inadequacies of these systems.

4. Discussion

In formal organism systems such as those we have been discussing here, there is much which probably will prove inapplicable to naturally occurring biological systems. The formal approach to biological theory however does allow us to deduce the precise consequences of postulating particular rules of growth, development, and behavior. Of course the formal representation may considerably distort a biological organism or process and so we must evaluate formal systems (such as our autonomous-cell systems) not only for their adequacy in producing the desired ultimate effects, but also for their directness and naturalness in representing phenomena of interest.

In showing that every organism form could be created employing only (restricted) autonomous-cell rules we surely did not thereby adequately "explain" the development of form in living systems generally. The scheme employed there of completely deterministic mechanical "unfolding" may be an adequate description of some phases of development, but it cannot be the only developmental process operative in organisms. For the system employed there obliges a single "master cell" to possess explicitly all the information on the creation, position, and type of every cell of the organism. Although this master cell creates new cells and cell connections, its action is in no way affected by the presence or absence of cells impinging on it. This means that the system (as there formulated) has no capability for self-repair; interaction with the larger environment also is not expressible.

In our example of organism universal *behavior* we did permit cells be sensitive to the presence or absence of impinging cells.

There is thus, in such a system (in contrast to the mechanical unfolding system first considered) some capability for self-repair and environmental interaction. In our universal behavior organism, we made use of this limited sensitivity to surroundings

to devise a technique for transferring information from one cell to another. The "crowding effect" employed is not likely to explain all inter-cell communication. Such a technique, in practical biological terms, would probably be very error prone, since it is not reasonable to expect that a cell could always detect on its surface the difference between 9 cells and 10, or 50 cells and 51. It is also biologically unreasonable to re-quire the propagating cell to produce large numbers of cells in a single act of multiple fission. Also, in order to make a cell again receptive to information from other cells, it must slough off all its appended signalling cells (which we can then assume to be lost). Transfer of information thus imposes a considerable metabolic cost on the organism.

All these limitations can be somewhat eased by taking advantage of the result by Shannon (1956) that only *two* Turing machine reading-head states are necessary (with a consequent increase in tape -symbols) to achieve universal computation. Thus, in our artificial computing organism we could have the appending of *one* cell mean state 1 is to be adopted and *two* appended cells mean that state 2 is to be adopted. (If we have any fears that a cell may not be able to distinguish between the addition of one cell or two, we could of course signal the contrasting states to be adopted by producing contrasting propagations of one appended cell versus some number of cells large enough to make the distinction clear. This will exacerbate again the multiple fission problem.)

We can of course have the propagating cell produce any required number of new cells by *successive binary* fissions. With this approach however we need to insure that the cell receiving the new cells does not begin to "interpret" the effect upon it until the propagating cell has completed its divisions (else the receiving cell may differentiate prematurely and behave incorrectly). One solution to this is to impose a (fixed, finite) internal lag or pause in taking action by any receiving cell long enough so that

the maximum propagating effect of a cell neighbor will have been completed.

There are other results established which permit us to re-design an organism so as to preserve organism ultimate capabilities while bringing the organism more into conformity with known biological reality. For example, it can be shown that our universal organism will retain its full behavioral powers even if we impose upon it the requirement that it may extend its length by adding new cells at *one end* only. Also, as in many other universal systems, the "halting problem" will be unsolvable for our organisms. That is, it will in general be impossible to predict ahead of time whether the growing and state-changing behavior of the organism will ever cease.

Although one of the principal advantages of a formal approach to biology is that we are able to draw upon the large body of known results and techniques in automaton theory to obtain new results in biological theory, we must here proceed very cautiously. For example, it is *not* true in our system that if we are given the developmental rules and given an artificial organism, putatively a consequence of the rules, it will be impossible to tell in general whether the organism *is* producible by those rules. In the system (as we presented it in the last section) the length of the given organism will enable a bound to be set on the number of times the rules might have been applied, and thus define the maximum length of an investigation to establish organism legitimacy or not. We *can* get the classic impossibility result if we introduce into our system rules which always detach any cells which are both *end* cells and whose tape-state is *blank*. In this case, the potential great expansion and contraction of the organism size would make setting a bound on the length of an investigation impossible, and the desired result would follow. Note however that we must assume that the detached cells will have been dispersed (along with the

sloughed-off signalling cells) or else a count of them might be employed to establish a bound on the past active life of the organism, and from this determine whether it could be a consequence of the rules.

This perhaps gives some indication of the complex ways in which formal methods and results on the one hand, and biological phenomena on the other, can interact.

Despite the inadequacies of one or another formal system in explaining biological phenomena, abstract automata, the capabilities of which are very conveniently expressed in the form of re-write rules, seem to provide an excellent source of models, and it may be very fruitful to view formal biology as a special branch or interpretation of automaton theory. Formal biology in this sense would then be seen as a theory of *all* organisms, both natural and artificial.

References

Laing, R. and Arbib, M. (1972) "Automata Theory and Development, II" The University of Michigan Technical Report. Submitted to *J. Theor. Biol.*

Minsky, M. (1961) "Recursive Unsolvability of Post's Problem of "Tag" and Other Topics in Theory of Turing Machines" *Annals of Math., 74,* 437-454.

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

Turing, A. (1936) "On Computable Numbers, with an Application to the Entscheidungsproblem" *Proc. London Math. Soc.,* Ser. 2-42, 230-265.