MEASURES OF ORGANIZATION IN A MODEL OF CELLULAR SELF-REPRODUCTION BASED ON TURING MACHINES

WALTER R. STAHL
OREGON REGIONAL PRIMATE RESEARCH CENTER
BEAVERTON, OREGON

1. Introductory summary

The late John von Neumann first realized that self-reproduction, as observed in biological systems, was a legitimate problem of mathematics, specifically of automata and algorithm theory. He founded the study of self-reproduction in systems of "natural automata." This report deals with the self-organizing properties of a new type of cellular self-reproduction model. The action of enzymes is simulated by Turing machines acting as molecular automata or computers, with their highly standardized coding corresponding to genes of the cell. Computer experiments have been done to test the stability and logical homeostatic properties of a 36 gene cell model. It is shown that the system will continue to organize itself and reproduce in spite of a variety of environmental deficiencies and disturbances, and also to repair itself after certain kinds of injury. The most pertinent organizational criteria for the described model are amounts of "energy" and "time cycles" needed to reach maturity and reproduce. Comparisons are drawn between these organizational measures and thermodynamic informational parameters such as Gibbsean chemical potential and entropy. The total amount of genetic and cellular information can be quantified in the cell model and this helps clarify some confusing aspects of genetic information measures. The model is highly idealized. It bears somewhat the same relationship to real cells as computer circuits do to the brain, but shows that automata theory can be applied to molecular biology in a meaningful way.

2. A model of self-reproduction based on Turing machines

Previous reports [1], [2], [3] described a cell model in which the principal logical tool is a molecular automaton that acts upon substances encoded as strings of symbols. The cell consists of a group of such enzyme automata that

Supported by Grants GM 13605-01 and FR 00163 of the National Institutes of Health. Publication No. 98 of the Oregon Regional Primate Research Center.

travel about a "cell contents tape," which contains all substances and structures of the cell in symbolic form. The molecular automata are coded in the notation of the Turing machine [4], [5], which is an absolutely standard format consisting of a quintuplet of elementary symbols. The Turing machine or automaton is a classical tool of algorithm theory [6], [7], [8], which has been used to prove a number of theorems in logic. Its use to simulate molecular automata is new and of interest because it makes it possible to represent cellular control algorithms in a manner compatible with modern automata theory.

The cell model consists of 36 enzyme automata with their "genetic coding." It is in no sense a quantitative simulation of cellular enzyme kinetics and is not intended to predict any biochemical facts. The system is a qualitative logical model designed to study the nature of control algorithms required for self-reproduction and self-organization. An excellent precedent for such highly idealized studies of biological units is the McCulloch-Pitts [9] "logical neuron," which is not an accurate model of any neuron, but made possible the design of "thinking networks." The "logical neuron" was well known to von Neumann and, in fact, played a rather important role in stimulating development of modern digital computers, first designed in about 1947–49.

Most reports dealing with "self-organizing systems" are based on the McCulloch-Pitts neuron, usually incorporated into a learning network which organizes itself in the sense of reacting purposefully to a certain kind of environment, that is, it learns to adapt to a given set of somewhat predictable environmental inputs. In the model of this paper, self-organization is of a different type, more comparable to cellular differentiation, in that the 36 molecular automata function cooperatively to produce a variety of complex products, cause the cell to reproduce, and adapt to environmental conditions.

3. Method of computer simulation

The cell model is run on a SDS-920 computer using a specially written compiler called TASP (Turing Automaton Simulation Program), which is fully described in a report by Coffin, Goheen and Stahl [10]. This simulator functions in the manner of a "universal Turing machine," as this term is used in the technical literature. It accepts the coding for any Turing machine (such as an enzyme automaton), together with some appropriate coding region upon which it is supposed to act (the cell contents), and simulates action of the Turing machine defined by the coding. The simulator operates at processing rates of 1000 to 5000 commands per second, has various necessary provision for editing, printing, error detection, and so forth. Direct printouts from TASP runs are reproduced in figures 4 to 6 below.

Because of its indirect mode of action, which is doubly removed from actual machine logic and circuitry, the TASP simulator is rather slow. Recently, therefore, another type of simulation program called CLPP (Cellular List Program Processing) was designed and speeds up operation by about a factor of 1000 on

the same machine. This compiler will be described in detail elsewhere, but is fully operational and now being used to study colonies of cells similar to the ones described in this report.

4. Operation of the cell

Details of operation of the self-reproducing cell are given in recent reports [3], [11] and will be reviewed only in brief. Figure 1 is a graphical representation of the cell model contents and represents the overall cell structure. Basic functioning of the cell model is exemplified by action of the enzyme automaton with a gene identification number 1411. This number is the greatly shortened surrogate for the 1000 or so *numerical* Turing machine code instructions that are actually used in the TASP simulator. It symbolizes the fact that the enzyme coding can be written in absolutely standard numerical form, accepted by a universal Turing machine. The enzyme represented (encoded) by gene program 1411 carries out the "string synthesizing" step: $W + A = SA$, in which W and A are elementary substances of the cell and SA is a more complex one, being the mnemonic for "sugar A." This step proceeds if, and only if, there are adequate levels of W and A, there is not too much of SA, there is enough energy E, and also is controlled by presence of a special controller string J. If all these threshold conditions are met, enzyme automaton 1411 produces six units of SA per time cycle, with corresponding losses of A, W and E.

All the enzymes function in this manner. The complete set of gene enzymes specifications is available elsewhere [3], but are shown in abbreviated form in table I. Figure 2 indicates the overall flow chart of the cell. The basic idea is to use elementary "diet letters" as the starting point for all cell products, which are generated through sequences of interlinked enzyme automaton steps. The cell makes the symbols 1, 2, 3, 4, which represent DNA coding (I, 2, 3, 4 for RNA coding) and A, B, C, D, which are the equivalents of amino acids. It also contains various biochemicals such as FA, FB, SA, and so forth, simulated organelles, such as OA, which stands for a mitochondrion and is needed to generate energy E, and OB, which functions as a ribosome. Chiaraviglio [12] has suggested that the ribosome can be regarded as a "sequential machine" (finite automaton) during its mRNA to amino acid transcription processing, and there is clear analogy to this in the cell model.

Operation of gene 1411 will be traced through to illustrate how the machine cell functions. The cellular gene enzyme automaton is identified in the TASP simulator as automaton program 1411 and activated once in each time cycle. Its first step is to find the simulated DNA region "D1411" and see if it is "on," as shown by a +. If this is the case, a complementary RNA is generated from the symbols I, 2, 3, 4, namely "R4I44", in the presence of a special enzyme acting as a RNA polymerase. If R4I44 and the ribosome automaton are present, an "enzyme string" $DADD$ is produced (all enzymes are coded in the A, B, C, D alphabet). These steps involve the coordinated action of the whole cell, which

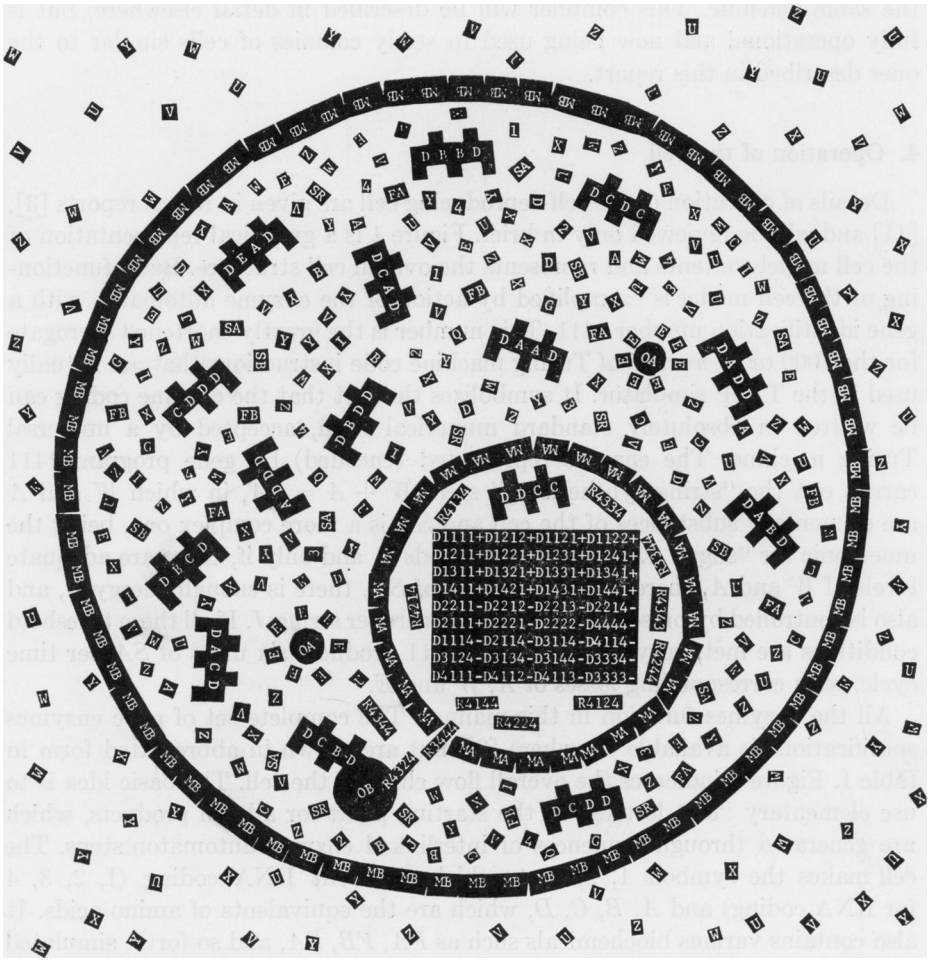

FIGURE 1

Graphic representation of the algorithmic cell model.
All substances are shown as letter strings,
including the bounding membranes of the cell and nucleus.
The genes are shown with an activation state symbol (+ or −).
Enzymes are represented by the complex geometric blocks
whose ends represent multiple tape reading sites.
Space limitations prevent showing the full number of
strings in the cell model.

generates the necessary symbols 1, I, 2, 3, 4, and A, B, C, D, as well as E, and so forth. If an enzyme has been generated symbolically in the above manner, then finally it can function within the cell, to produce SA from W and A, under the stated threshold conditions.

TABLE I

SUMMARY OF GENE ENZYME AUTOMATA ACTIONS

A complete table with rates and threshold conditions is given in Stahl [11].
A complementary copy is made of DNA number in RNA symbols; the enzyme matches the RNA number and is made by action of a ribosome automaton; the numbers and letter sequences are surrogates for much longer coding regions defining the Turing automata (which are too long to show in the cell model).
Asterisk indicates enzyme not produced by these genes, whose RNA is a regulator or controller.
Dagger indicates distribution to two daughters is slightly randomized.
Later models of this system have a more detailed simulation of cell wall formation, organelle distribution, and so forth.

DNA Gene	RNA Copy	Enzyme	Action
1111	4444	*DDDD*	$Z + Y = E + G$ (energy generation)
1212	4343	*DCDC*	$V + Z =$ I (RNA nucleotide)
1211	4344	*DCDD*	$U + V =$ 1 (DNA and RNA nucleotide)
1221	4334	*DCCD*	$U + W =$ 2 (DNA and RNA nucleotide)
1231	4324	*DCBD*	$U + X =$ 3 (DNA and RNA nucleotide)
1241	43I4	*DCAD*	$U + Y =$ 4 (DNA and RNA nucleotide)
1311	4244	*DBDD*	$V + Z = A$ (amino acid)
1321	4234	*DBCD*	$V + W = B$ (amino acid)
1331	4224	*DBBD*	$V + X = C$ (amino acid)
1341	42I4	*DBAD*	$V + Y = D$ (amino acid)
1411	4I44	*DADD*	$W + A = SA$ (sugar A)
1421	4I34	*DACD*	$W + B = SB$ (sugar B)
1431	4I24	*DABD*	$Y + A = FA$ (fat A)
1441	4II4	*DAAD*	$Y + B = FB$ (fat B)
2211	3344	*CCDD*	$SA + FA = MA$ (membrane A)
2212	3343	*CCDC*	$SB + FB = MB$ (membrane B)
2213	3342	*CCDB*	$MA + FA + 1 = OA$ (mitochondrion)
2214	334I	*CCDA*	$MB + FB + 2 = OB$ (ribosome)
2222	3333	*CCCC*	$I + 2 + 3 = SR$ (soluble RNA)
1114	444I	*DDDA*	$I + 1 + 4 + A = OT$ (test automaton)
1121	4434	*DDCD*	$W + X = U$ (alternate source U)
2111	3444	*CDDD*	$D +$ I $= J$ (maturation controller)
2221	3334	*CCCD*	$C + MA + 3 = Q$ (reproduction initiator)
1122	4433	*DDCC*	RNA polymerase action
2114	344I	—	maturation sequence status (via J)*
3114	244I	—	checks premitosis ready status (via Q)*
3124	243I	—	premitosis check step one*
3134	242I	—	premitosis check step two*
3144	24II	—	mitosis possible*
3334	222I	—	deactivate maturation genes*
4114	I44I	—	deactivate long string genes*
4111	I444	*ADDD*	DNA polymerase action
4112	I443	*ADDC*	split all reservoir in two†
4113	I442	*ADDB*	separate gene strings (print)
4114	I44I	—	activate general growth genes*

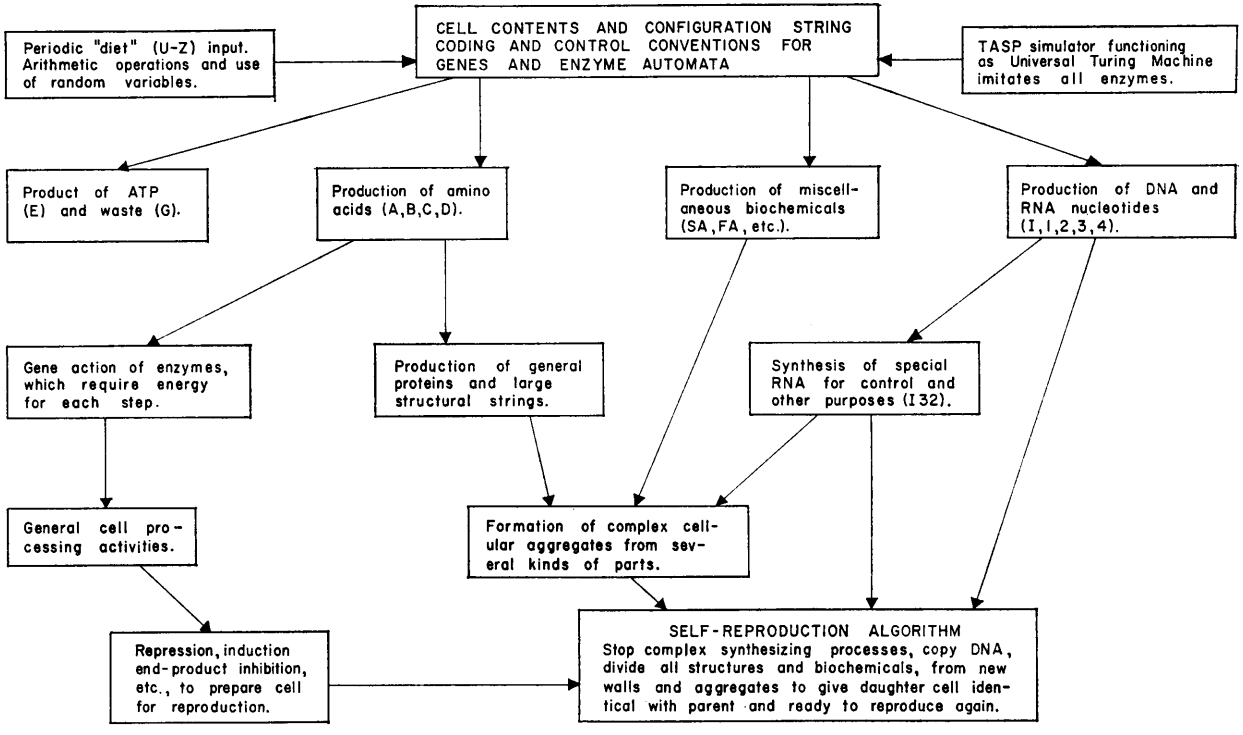

FIGURE 2

Computational flow chart for the string processing self-reproducing system.

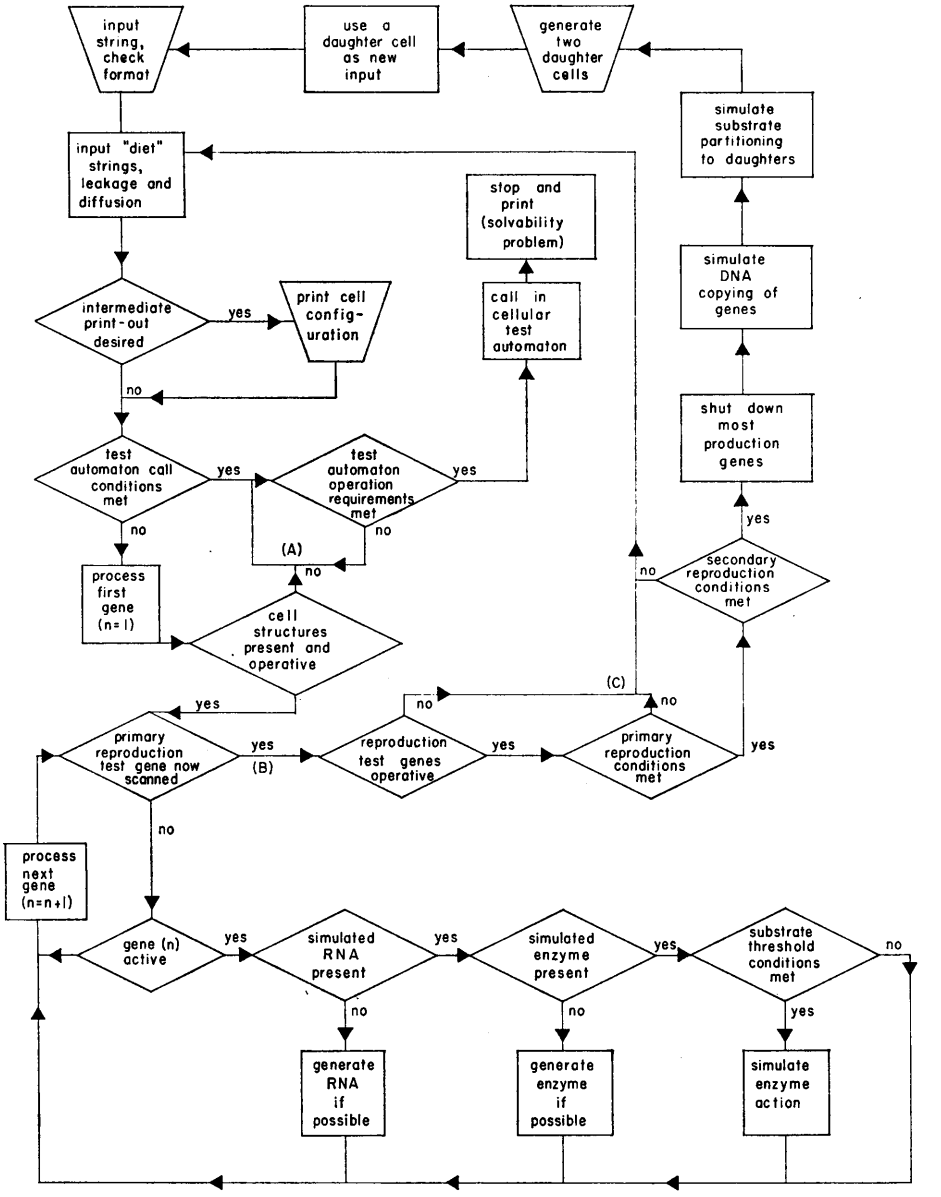

FIGURE 3

Computational flow chart for the cell model.

Figure 3 is the computational flow chart for the cell model. Each gene enzyme automaton is activated once per time cycle of operation, which is sufficiently short to simulate parallel operation of the gene enzyme molecular automata. There are a number of heuristic provisions for error stops; changes of substance

levels during test runs; keeping counts of total use of energy and total time cycles of operation; printing output configurations of the cell, and so forth, which shall not be discussed further in this report. Table I shows the "string synthesizing" and other actions of the 36 cell gene enzyme automata.

5. The self-reproduction algorithm

The general course of self-reproduction is evident from figures 3 to 6. At birth certain basic genes for production of energy, RNA polymerase, the RNA and protein symbols, and so forth, are operational in the cell. There is a progressive buildup of longer strings and generation of special controller strings, such as Q or J, which turn on genes that were not active at the outset. The change in pattern of positive genes is very obvious between figures 4 and 5. The self-repro-

```
(1)NORMAL BUILD UP IN PROGRESS          C=/5
/1111+RE  /1212+RE  /1121+RE  /1122+RE  /
/1211+RE  /1221+RE  /1231+RE  /1241+RE  /
/1311+RE  /1321+RE  /1331+RE  /1341+RE  /
/1411+R   /1421+R   /1431+R   /1441+R   /
/2211-    /2212-    /2213-    /2214-    /
/2111+R   /2221-    /2222-    /4444-    /
/1114-    /2114-    /3114-    /4114-    /
/3124-    /3134-    /3144-    /3334-    /
/4111-    /4112-    /4113-    /3333-    /
S/E=009//G=30/
/U=45//V=53//W=73//X=64//Y=95//Z=94/
/1=68//2=64//3=59//4=12//1=40/
/A=07//B=00//C=00//D=05/
/SA=00//SB=00//FA=00//FB=00/
/MA=02//MB=02//0A=02//0B=04/
/J=00//Q=00//SR=02//0T=00/
```

FIGURE 4

Computer printout of cell configuration
during an early stage of processing.
The comment on top line is an heuristic for
use of operator and not part of the cell algorithm.

duction algorithm is basically a program to synthesize all the cell symbols, including the gene (Turing machine) coding numbers, which are taken to exist in the cell in the manner of "alphabet soup." After accumulation of sufficient symbols specific enzymes simulate action of DNA polymerase to put together another copy of the gene coding actually using symbols 1, 2, 3, 4—this repre-

sents copying of the DNA chain. Noncomplementary direct transcription is used for purposes of convenience, but double stranded complementary coding could be simulated easily. The cell model thus replicates the numerical coding programs of all the cellular automata in a standard manner.

Figure 6 shows a "daughter cell" as it actually is produced by the TASP

```
 (7) DAUGHTER  2                                              C = /3
/1111 + RE  /1212 + RE  /1121 + RE  /1122 + RE  /
/1211 + RE  /1221 + RE  /1231 + RE  /1241 + RE  /
/1311 + RE  /1321 + RE  /1331 + RE  /1341 + RE  /
/1411 -     /1421 -     /1431 -     /1441 -     /
/2211 -     /2212 -     /2213 -     /2214 -     /
/2111 -     /2221 -     /2222 -     /4444 -     /
/1114 +     /2114 -     /3114 -     /4114 -     /
/3124 -     /3134 -     /3144 -     /3334 -     /
/4111 -     /4112 -     /4113 + RE  /3333 -     /
S/E=024//G=40/
/U=20//V=30//W=22//X=24//Y=22//Z=20/
/1=30//2=12//3=04//4=00//I=16/
/A=14//B=18//C=20//D=06/
/SA=06//SB=06//FA=06//FB=06/
/MA=06//MB=04//0A=04//0B=04/
/J=06//Q=04//SR=04//QT=00/
```

FIGURE 5

A number of substances have been built up and genes turned on; a sequence of gene tests of reproduction conditions have been met.

system. This configuration can be recycled automatically in the simulator and will continue to reproduce itself provided only that it is supplied with a sufficient quantity of "diet letters" within a certain time period; if it is not, the cell reaches a point of irreversible internal insufficiency and "dies." The action of the cell model is not rigidly mechanical and in any given time cycle there is no obligatory completion of any fixed number of steps. The "self-organizing" nature of this process will be discussed below.

The basic strategy of the cell program, which represents a new concept of artificial self-reproduction proposed by the author, is as follows: (1) everything in the cell is represented by symbols on a long tape; (2) all substances are built up by "string synthesis" from elementary "diet letters" supplied from the outside; (3) the cellular automata move within the cell contents space and synthesize products with "conservation of letters"; (4) the cellular synthetic automata function at fixed rates but have multiple threshold determinants for substrates, products and controller strings; (5) energy (the string E) is required for any

enzyme to function (but not for threshold testing) and must also be generated from the diet; (6) special automata function as mitochondria, ribosomes, RNA or DNA polymerase, and even simulate "active transport" or diffusion through the cell walls; (7) self-reproduction consists in building up all the necessary

```
(3) REPRODUCTION POSSIBLE                                    C=1 /
/1111+RE  /1212+RE  /1121+RE  /1122+RE  /
/1211+RE  /1221+RE  /1231+RE  /1241+RE  /
/1311+RE  /1321+RE  /1331+RE  /1341+RE  /
/1411+RE  /1421+RE  /1431+RE  /1441+RE  /
/2211+RE  /2212+RE  /2213+RE  /2214+RE  /
/2111+RE  /2221+RE  /2222+RE  /4444-    /
/1114+    /2114+R   /3114+R   /4114-    /
/3124+R   /3134+R   /3144+R   /3334-    /
/4111+    /4112-    /4113-    /3333-    /
S/E=082//G=50/
/U=40//V=70//W=46//X=50//Y=94//Z=80/
/1=84//2=62//3=42//4=99//I=46/
/A=31//B=43//C=42//D=40/
/SA=14//SB=12//FA=14//FB=14/
/MA=12//MB=10//0A=08//0B=08/
/J=12//Q=08//SR=08//0T=00/
```

FIGURE 6

This shows one of two new "daughter cells" printed by the TASP system. They need not be quite identical in quantities of strings, but are both in the required starting gene configuration. Comment on top line is for use of operator and not part of cell model proper.

symbols to make a new cell configuration, as shown in figure 6, with template copying of the gene coding and splitting up of reservoirs of all strings in the cytoplasm.

Gene (DNA) reproduction is symbolized by copying of the four place gene enzyme identification numbers; the actual Turing machine coding can easily be written in the (1, 2, 3, 4) coding, but would, of course, be much longer than the four place gene identification number. In principle, however, copying the identification number at the proper phase of cell development shows that the whole Turing machine coding for each gene enzyme automaton could have been copied at this juncture in the reproduction algorithm. A considerably larger cell, but not one that is logically more complex, would be needed to actually copy the entire automata codes.

Probably the most important single idea in the whole model is that the gene program codes are not "written" but assembled out of symbols that have been made synthetically, and are floating in the cell as in alphabet soup. No other known model of self-reproduction uses this approach to construction of a new program code in "empty space," defined by indefinitely extended tape of a Turing machine (or machines). It is conceptually related to the von Neumann kinematic model which was to make a new physical machine in a "sea of parts." His tesselation model had a highly structured space in which new configurations were to be reproduced. The described system is fully operational on a computer; insofar as is known, the other von Neumann models have not yet been fully simulated on a computer.

6. Relationship of model to biological cells

Table II compares the contents of the cell model with a real cell. The validity of such a highly abstract model is discussed much more fully in other reports [2], [3]. It is never possible to prove a model is "true" and models are frequently most valuable in an heuristic, conceptual sense.

The described model does not deal with quantitative biochemical phenomena of cells, and in this regard is completely different from the enzyme kinetics models described by Chance, Higgins and Garfinkel [13], Garfinkel and Hess [14], and others. It does include a great many qualitative relationships and control mechanisms that are presently believed to operate in cells, as described in books on molecular biology by Bonner [15], Watson [16], Morowitz [17], F. W. Stahl [18], or Paul [19]. The model specifically incorporates the principles of DNA regulation by RNA suggested in the work of Jacob and Monod [20], and Changeux [21], various biochemical feedback mechanisms cited by Eagle [22], or Morowitz [23], and represents DNA in the basic roles described by Commoner [24], Zuckerkandl and Pauling [25] and others.

Recent surveys of progress in exobiology by Pattee [26], [27], Morowitz [17], and Bernal [28], have made it increasingly clear that cellular self-reproduction should be regarded as a general process, not necessarily dependent entirely on availability of materials on the earth such as DNA. In a general sense the cell model described in this report is comparable to the PPLO organism, which has been discussed extensively by Morowitz and Tourtellotte [29]; figure 1 even looks like the illustration these authors gave for the PPLO organism.

An important measure of organizational complexity of a cell is its total content of molecules, macromolecules and genes. A thorough analysis of the PPLO (which is probably the smallest organism with a membrane) by Morowitz [30], [17], reveals that even this very simple structure includes about 10^8 to 10^9 molecules other than water molecules (compare letters in the model), about 10^7 monomer units (DNA, RNA protein coding letters in the model), some 18,000 large molecules (longer strings of the model), and 400 enzymes or proteins. Hutchinson [31] calculates that a normal cell has 10^{14} atoms of hydrogen and

TABLE II

A COMPARISON OF CODING AND CONTROL MECHANISMS IN THE STRING PROCESSING MODEL AND REAL CELLS

	Natural Cell	Automaton Cell Model
Coding	Four letter DNA code, with about 1000–2000 nucleotides per gene; 100,000 or more genes per cell.	Turing quintuplet code, with automata identified by "gene numbers." About 1000 Turing commands per enzyme automaton and 36–80 automata per cell.
Copying	DNA subcopied via four letter mRNA, under operator, repressor, *etc.*, control. Direct copying by DNA polymerase.	Automata (gene) coding copied symbolically via RNA subcopy strings. Actual copying is from tape into core memory of computer, once for each gene during each processing (time) cycle.
Enzyme formation	Chains of mRNA serve as templates for ribosomes, which act as universal polypeptide constructional machines.	The Turing code of a particular enzyme automaton is processed in computer, acting as universal Turing machine, which simulates each enzyme in turn.
Enzyme action	Cell synthesis and logical activity controlled by enzymes which can recognize and operate on all biochemicals and cell structures.	All biochemicals and cell properties represented as symbols on program tape. Enzyme automata can perform any desired operation on these symbols. No other means of manipulating them is used.
Structural features	Proteins and other biochemicals build up complex cell structures such as membranes and organelles.	Automata combine and transform various biochemicals to give complex two dimensional arrays or structures, which can polymerize.
Redundancy	Many copies of genes, enzymes, ribosomes, all biochemicals, cell wall elements, *etc.* Much error tolerated, except in certain key genes.	There is one copy of basic coding, but cell is operated as if there are many copies of RNA, enzymes, proteins, biochemicals, *etc.* Considerable error tolerated in rates and some aspects of coding.

oxygen, 10^{12} to 10^{13} of carbon and nitrogen, 10^{10} to 10^{12} of phosphorous, sulphur and a number of other elements, and even 10^{4} to 10^{6} molecules of mercury, uranium, and so forth, present as trace contaminants (these correspond rather loosely to certain heuristic programming symbols used in the TASP program). The automaton cell does not approach this size, but it could perhaps be made to do so within the foreseeable future. Even so, it takes many millions of elementary logical enzyme automaton steps to mature.

The described cell simulation is a rather more plausible model of biological cellular self-reproduction than the so called "tesselation model" of von Neumann [32], as it has been developed in the hands of Moore [33], Burks [34], Myhill [35], Arbib [36], [37] or Barzdin [38]. In the tesselation model there is an indefinitely large grid of squares, each of which contains a completely prepro-

grammed identical specific finite automaton, which is quiescent until activated by adjacent cells. Reproduction in this model is of a configuration in the grid, and not of the automata in each cell of the grid as such. Arbib [36], [37] has recently suggested some very interesting variants on the tesselation model in which the program code of a Turing machine can be self-reproduced into adjacent cells, but the surrounding region is by no means "empty" and in fact highly structured. The model described in Stahl [3] is the only known automata theory model of a cell which generates a new cell next to itself, from elementary parts, repeatedly, and also the only such known model which attempts to use known mechanisms of molecular biological control. In some respects it is a (molecular biological version) of von Neumann's original kinematic model, which was intended to reproduce an actual physical array of transistorlike parts in space. It does, of course, have its own deficiencies, connected principally with quite arbitrary definitions of environment, enzymes, genes, and so forth.

7. Testing of self-organizing ability of the cell model

Tables III, IV and V deal with computer experiments in which the normal cell model, which reproduces as described above, was subjected to various kinds of interference. Normal maturation time with the TASP compiler was 2 to 3 hours, and only 5 to 10 minutes with the new CLPP system. Both compilers have a provision for interrupting cell operation (at the end of any processing cycle), with arbitrary readjustment of levels of any substances, rates, activation condition of genes, and so forth, and also for applying perturbations at the end of every normal operational cycle (continuing disturbances).

A large variety of numerical parameters can be extracted during operation of the cell model, such as total use of any "diet" string, turnover rates of certain strings, relative ratios of strings, and so forth, but the following seemed the most natural and significant indicators of cellular self-organizing power: (1) number of operational time cycles to reach maturity (as signaled by appearance of a certain control string J); (2) number of operational cycles to reproduce in accordance with the full normal self-reproduction algorithm, and (3) total energy E expenditure for reproduction (it will be recalled that E is made by an enzyme automaton from elementary diet strings). More complex criteria are of value for certain tests, but will not be discussed here.

Table III shows the results of studies in which a certain fraction of the total quantity of all small substances (one and two letter strings) was removed from the cell at the end of each time cycle. This corresponds to "leaky" cell walls, with loss in proportion to total quantity, $dN/dt = kN$. Table III shows that the cell could reproduce with loss rates up to about 30 per cent, and simply survive with considerably higher leakage rates (up to at least 50 per cent, though this is not shown in table III). Loss rates up to 8 per cent had little effect on time and total energy to reproduction, while values of 16 to 30 per cent caused considerable delay. Maturation with production of large cell strings and organ-

TABLE III

TESTS OF THE ORGANIZING ABILITY OF A CELLULAR SELF-ORGANIZING MODEL ON TURING AUTOMATA

Effects of Leakage of All Small Substrates
Gene enzyme automaton model with TASP simulator (9/65).
Leakage of all one and two letter strings, at end of each time cycle, at percentage of their total contents shown.
NR indicates no reproduction in 1000 time cycles; it will probably never take place.

Leakage Rate (per cycle in per cent)	Organizational Criteria		
	Cycles to Mature	Cycles to Reproduce	Total Energy to Reproduce (rounded to units of 50)
none	35	46	1550
2	35	46	1600
4	38	60	1700
8	66	75	1750
16	72	87	2000
25	72	100	3300
28	76	121	4000
30	89	360	11900
32	89	NR	NR

elles was not nearly so demanding a self-organizing task as reproduction, which is to be expected.

Table IV illustrates comparable testing in the situation that certain specific "diet" substances were in short supply. Dropping the input available on each processing cycle of the diet strings W, U and V to 25 had very little effect on the standard organizational measures. Decrease of Z or Y caused more difficulty

TABLE IV

TESTS OF THE ORGANIZING ABILITY OF A CELLULAR SELF-ORGANIZING MODEL BASED ON TURING AUTOMATA

Effects of Environmental Deficiencies
Thirty six gene enzyme automaton model with TASP simulator (9/65).
The normal outside letter diet supplied each cycle is curtailed in manner shown.
NM indicates no maturation and NR, no reproduction in 1000 cycles.

Nature of Shortage	Organizational Criteria		
	Cycles to Mature	Cycles to Reproduce	Energy to Reproduce (rounded to units of 50)
none	35	46	1550
25% normal W	38	56	1550
25% normal U	53	59	1700
25% normal V	87	98	1600
12.5% normal Z	203	338	16900
12.5% normal Y	220	358	17900
6.2% normal Z	NM	NR	NR

and inputs of 12.5 per cent of normal for these materials caused a major slowdown in reproduction, which, however, was completed in due course, with a normal daughter cell configuration, but with use of over 10 times the normal energy amount. Reduction of the substance Z to 6.2 per cent of normal environmental values (at each operational cycle) was beyond the organizing limit of the cell: it did not die, in the sense that its E level dropped finally to zero, but was not able to mature.

Table V shows the results of a different sort of testing, in which during normal

TABLE V

TESTS OF THE ORGANIZING ABILITY OF A CELLULAR SELF-ORGANIZING MODEL BASED ON TURING AUTOMATA

Effects of Abrupt Reduction to Zero of Certain Substances
Thirty six gene enzyme automaton model with TASP simulator (8/65).
Normal operation in optimum environment, without leakage,
and some materials are suddenly ablated (dropped to zero).
NR indicates no reproduction in 1000 operating cycles;
self-reproduction algorithm not viable.

	Organizational Criteria		
Nature of Disturbance	Cycles to Mature	Cycles to Reproduce	Energy to Reproduce (rounded up in units of 50)
Normal	38	47	1600
Destroy RNA enz. of genes 1411, 1421, 1431, 1441, at 17th cycle then proceed	43	51	1700
Ablate *J*, *D*, and I on 18th cycle and proceed	40	52	1700
Ablate reservoirs of I, 2, 3, and *SR* on 38th cycle and proceed	43	53	1800
Ablate reservoirs of *Q*, *A*, *RA*, *MA*, *SA*, and *C* on 39th cycle and proceed	41	53	1800
Turn off permanently gene 1122 or 2114 or 3114 or 3334 at 10th cycle	47–200	NR	NR

maturation there was a sudden complete removal (ablation) of one or more substances, with subsequent continuation of the test run. It shows that the cell would recover easily from loss of certain of its simulated equivalents to messenger RNA or enzymes, simply by making more copies of these materials. The cell system would also recover easily after dropping of most string reservoirs to zero at any arbitrary processing stage. This sort of action is possible because a wide variety of interlocking feedback loops, similar to end product inhibition of allosteric enzymes, and Jacob and Monod RNA regulation mechanisms, are built into the cell control algorithm [11]. Multiple threshold tests for the enzymes cause them to maintain the proper levels of their principal products rather independently. Sudden loss of a reservoir results in a propagated release of

thresholds and rather rapid resynthesis of needed strings, at which time the basic reproduction algorithm proceeds.

The last line of table V deals with disturbances which cannot be restored by the cell. Turning off of a gene at an arbitrary time may result in configuration which the cell cannot reverse. For example, if the normal maturation sequence calls for activation of gene no. 2114 at the end of cycle 8, but it is then arbitrarily deactivated on cycle 10, the cell algorithm is unable to reorganize itself and never again returns to the configuration recognized as appropriate for turning on gene 2114.

This response is somewhat reminiscent of, but not identical with, the "Garden of Eden" configurations described by Moore [33] in his studies of von Neumann's tesselation models. Such configurations cannot be reproduced and are forever eliminated in Moore's system. In the described model recovery is possible following deactivation of some genes, particularly those involved in premitotic testing, but not most of the important controller genes, which indicate general level of cell maturation. Interference with the genes involved specifically with reproduction during the mitotic process itself usually has disastrous effects.

The self-repair capacity of self-reproduction models of the von Neumann tesselation type has been studied fully by Lofgren [39], who finds that certain kinds of error in the "grid" can be restituted by a given self-reproduction algorithm, which is programmed identically in all cell automata. The problem of self-repair through multiplexing of information channels was first clearly raised by von Neumann [40]. More recently, Winograd and Cowan [41] and Cowan [42] have increased the error resistance or self-organizing ability of neuron nets by more subtle types of computational redundancy. Much the same concept is present in the algorithmic cell design, in which each enzyme automaton is assumed to be present in a small number of multiple copies and "computes" a string synthesis which may influence a variety of other automata and gene control conditions. The cell model has considerable capacity for self-repair; models under study now include provision for lysis and elimination of enzyme automata, or whole cells which do not appear to be functioning properly, with replacement by new units.

An entirely different self-organizational adaptive problem is considered in the reports [11] dealing with what the author has called algorithmically unsolvable problems for the cell model. To demonstrate this result a "test automaton" is included in the gene program of the cell and has the hypothetical task of "computing" the composition of a new gene needed by it to fulfill a certain adaptive need. It can then be shown that for any well defined cell algorithm programmed in a standard manner one can always find a test situation or cell configuration which prevents effective adaptive action by the computational automaton. This result implies that direct genetic adaptation does not occur in nature because it involves fundamental logical paradoxes or dilemmas. The entire question is discussed in detail in the cited recent report.

From tables III through V, it is clear that the algorithmic cell model has

potent self-organizing capacities and will reproduce itself in spite of a wide variety of individual and combined disturbances. It does, of course, have its limits and sufficiently great losses or arbitrary interferences with certain substances or genes causes it to die, in the sense of entering into a progressive decline of basic substances and energy.

8. Other models of biological self-organization

Study of available symposia volumes on self-organizing systems edited by Yovits and Cameron [43], Yovits, Jacobi and Goldstein [44], Bernard and Kare [45] and von Foerster and Zopf [46], and also a recent volume in Russian by V. M. Glushkov [47], reveals that the self-organizing model described above is new, and that most prior work on self-organization has been concerned with learning in adaptive neural networks. The described model does not ever "learn the organization" of its environment, though its specific gene activation state can reflect the "diet" available in the environment. The specific difference between the cell model and neuron models is that threshold levels for the enzymes are never altered during normal cell operation in the former, whereas this step is the crux of learning and adaptation in almost all neural network models, the vast majority of which are based on the McCulloch-Pitts [9] "threshold logical neuron."

The described model also differs greatly from certain "automata games" proposed by Tsetlin [48], Gel'fand, Pyatetskii-Shapiro and Tsetlin [49], Borodyansky [50] and other Soviet authors interested in biological cybernetics. These authors are concerned with much smaller sets of automata, which are generally conventional finite automata existing in perhaps a dozen states and connected by binary signal channels. Working with Glushkov, Letichevskiy and Dorodnitsyna [51] developed a model of an adaptive automaton that feeds on binary bits scattered on its environment tape; this automaton must learn the presentation pattern of 1's on its tape or die of starvation, but it does not interact with other automata and there is no cell as such.

Important theoretical discussions of the problem of self-reproduction have been presented by Ashby [52], [53], Mesarovic [54], Muses [55], and von Foerster [56]. These authors treat self-organization as an abstract problem, but commonly refer to a learning network as the prime example of biological self-organization. Generally speaking they draw attention to what coded representation can be made of the environment, what takes place in the system during learning, and how its internal organization reflects the structure of its surrounding environment. All of these concepts are applicable to the algorithmic cell model of this report, but in a sense somewhat different from that appropriate to memory networks. Andrew [57] has discussed organizational systems in terms of their ability to maintain a stable external environment (as by a robot), but this task is rather different from that of maintaining a stable molecular *milieu interior* in the cell.

The described cellular model also differs in principle from those of evolutionary adaptation described in recent reports of Bremermann [58]. This author has provided interesting and detailed analysis of the informational selective processes implied by natural selection of a genome. He does not, however, propose a self-reproducing model as such, and self-organization in his type of system is treated from the standpoint of information theory.

In a series of fascinating and provocative papers McCulloch [59], [60], [61] discusses biological computers in a broad context and proposes the use of elementary logical tools for representing biological events. This same view was strongly emphasized by von Neumann [32]. The work of the author is the first attempt to use a large and complex system of Turing automata to simulate a biological cell engaged in molecular processing; it is also the first mechanical implementation of a universal Turing machine, which is the proper compiler of Turing machine codes, on an ordinary computer.

The model incorporates a number of complex biochemical regulatory mechanisms. The cell has been considered as a complex cybernetic regulatory system in recent reports of von Foerster [62], Beer [63], Apter and Wolpert [64], Frank [65], Dechev and Moscona [66], Aleksandrov [67], Kafiani [68], David [69] and others.

A large number of recent reports in molecular biology contain the terms feedback and molecular regulation in their titles, which illustrates the trend in molecular biology towards a mathematical cybernetic viewpoint. To the present there has been no agreement on what sort of model will be most useful for application of mathematical control and automata theory to molecular genetics. The model of this report is a possible vehicle for merging automata and control theory with molecular biology, but a number of others doubtless will be developed and studied in the near future.

9. Entropy and statistical measures of self-organization

In his small but provocative book *What Is Life?*, Schrodinger [70] some time ago raised questions about the relationships of statistical processes in large molecules, as viewed by quantum mechanics, and the organization of life. The applicability of quantum mechanical and statistical mechanical principles to cellular organization to cells have been discussed in reports by Elsasser [71], [72], Rosen [73], Longuet-Higgins [74], Karreman [75], and others, without formation of a consensus. Goodwin [76] provides a statistical mechanical approach to cell enzyme kinetics. The relationships between entropy and evolution or natural selection are considered in reports of Trincher [77], [78] and Ursul [79], without reference to a specific cell model or automata theory. It is of interest to attempt to compare the statistical organization measures of thermodynamics with those given above for analysis of organizational tasks in the cell model.

The first law of thermodynamics states that $E + H = W$, namely, total

energy is conserved but partitioned into heat energy and work. In the cell model there is a total quantity of strings or available energy E at any one moment, but the correspondence to heat and work energy is not immediately obvious. Temperature should be a factor causing an increased collision rate between all strings and hence augmenting all string reaction processes. Increase in temperature is expected to result in an increased content of heat (assuming a constant specific heat). In the cell model one might suppose that all strings are moving about randomly and have a kinetic energy proportional to size (number of letters) times their velocity. The equivalent of temperature will then be a mean linear velocity of motion within the cell. As the cell is actually modeled, what is moving is not strings but the string processing automata, which travel about the cell contents tape at finite "stepping rates." Random motion is represented in a similar manner, by an automaton adding random variables to all strings at a rate which depends on the postulated model temperature.

A good model of physical work development in a biological system is found in muscle, which is a structure that forces randomly oriented chemical processes to proceed along a particular axis. That is, regardless of the direction of motion of an ATP molecule when it arrives at a muscle fiber, this macromolecular machine releases the ATP energy along its fiber axes, which slide back and forth. No exact analogue to this mode of action is at present found in the cell model, but it could be added. An alternative generation of work is by a piston which forces random molecular motion to act in a particular direction. Work may be said to be an energy vector directed in a geometric physical or abstract phase space.

The second law of thermodynamics states that entropy ΔS increases with time in all physical processes into which there is no influx of "organizing energy." Entropy is commonly defined by $\Delta S = \Delta q/\theta$, with Δq being reversible heat content and θ temperature, and has the dimensions of heat content per degree. From the above definition of temperature, for the algorithmic cell model increase of entropy would correspond to an increasing ability of strings to move randomly, or with more degrees of freedom, thereby giving a higher total movement energy for a given cell configuration. This could be interpreted to mean that with increased temperature more strings will become free and not bound in aggregates or complex structures or pathways in phase space. Polymerization of "strings" was represented in a prior string processing cell model [80], and disruption of organized aggregates of strings appears as a natural result of increasing rate of string motion in this system.

The third law of thermodynamics states that entropy will be zero at absolute zero, which could correspond easily to complete cessation of all motion in the model. In figure 1 this would refer to a stopping of all string motion and contact of strings with enzyme automata, while in the actual TASP simulation it implies a dead stop of the computer, with freezing of a given configuration. It is of much interest that the cell model can be restarted after such a stop. Recently, Skoultchi and Morowitz [81] showed that cysts of the invertebrate

organism Artemia could be brought to absolute zero and then revivified, demonstrating that all genetic information necessary for survival of this organism was carried in its structure, and not in the pattern of electronic or molecular momentum distributions. The cell model of this report exists as a frozen cell algorithm until placed in the TASP computer simulator.

Entropy may also be defined by an equation of the form $-\Delta S = R \log (p_2/p_1)$, in which p_2 and p_1 are probabilities of a certain configuration or "state" and R is the gas constant. This equation can be interpreted to mean that entropy is *decreased* $(-\Delta S)$ when probability of a given configuration (among all possible ones) is increased. In the cell model increased orderliness is associated with regularities in enzyme automata configuration charts or the appearance of the cell (see figure 1), particularly in connection with aggregation of strings in certain regions at appropriate times, and formation of complex aggregates. In the model the string E is used for all logical work, and energy has to be expended to form any pattern (make its occurrence more probable in a statistical sense for the cell model).

Another very important concept of thermodynamics is that of chemical free energy ΔF or Gibbsean chemical potential ΔG. Defining equations for chemical free energy include

$$\Delta G = \Delta H - \theta \Delta S, \tag{9.1}$$

$$\Delta G = R\theta \log K_m, \tag{9.2}$$

$$\Delta G = R\theta \log (C_2/C_1) \tag{9.3}$$

in which ΔH is total heat energy, R is the gas constant in units of cal/mole deg, θ is temperature, K_m is a mass action constant representing a ratio of probability of formation over probability of destruction of a compound, C_2 and C_1, a higher and lower concentration of a given substance on two sides of some sort of barrier.

The first equation (9.1) can be interpreted in the model to mean that the total energy of effective organized motion (in a phase space) of the strings is given by their total motion less purely random motion; this has a reasonable counterpart in figure 1 and the automaton system. The second equation (9.2) states that free energy is required to shift a mass action constant in the direction of a desired product. In automaton cell model every enzyme makes product by simply using one unit of E per "bond," so concentration of product is linearly related to energy use, but a logarithm function will arise if a physiochemical randomizing automaton simulates breakage of bonds at a fixed probability for all bonds, as a function of temperature.

There is also a reasonable analogue for (9.3) in the model. In the latter work must be done, and the string E expended, to concentrate a particular string in a given region (or phase space), but there is no obligatory occurrence of the logarithm function relating energy use to a string concentration ratio. Such a function may arise if the randomizing automata carries out a passive physiochemical disruptive action act on all strings (particularly in the high concentra-

tion C_2 region), while the concentrating automaton moves one unit of C_2 across the barrier per unit of E, as done in the model ordinarily.

These examples show that provocative analogies to the concepts of chemical thermodynamics may be found in the cell model. However, total energy or time cycles for maturation or reproduction do not have an immediate counterpart in the basic measures of thermodynamics. But it has never been shown that negative entropy or free energy really are appropriate or the best organizational measures for biological cells, which are much more highly organized than any regular solid such as a crystal. Thermodynamic measures are not even particularly appropriate for transistors, and definitely not suitable for analysis of miniaturized solid state circuit or computer assemblies, which also have high levels of organization. They do not apply at all naturally to computers and neuron net models of organization.

10. Measures of information in the cell model

Possible applications of information theory in biology were discussed in some detail in a symposium edited by Yockey, Platzman and Quastler [82]. Information theory analyses have been made of the base sequences in DNA by Apter and Wolpert [64] and Gatlin [83], and of evolution by Bremermann [58]. An informational interpretation of aging has been given by Johnson [84], with the assumption that each cell contains a great deal of redundancy. Many works are available on informational processes in the nervous system, at membranes, during human communication, and so forth, but will not be cited. The cell model of this report makes it possible to explore some basic problems in the interpretation of genetic information content.

A Turing machine may always be coded in binary notation (as can any computer or robot), and this allows one to ascribe an information content to a region of cell coding. The information content of the model cell can be estimated from the following facts. It contains 36 genes with about 1000 Turing machine commands per gene enzyme automaton; a command consists of a quintuplet of symbols with a mean information content of about 24 bits. This makes a total of roughly 864,000 coding bits. The cell contents tape consists of some 1400 symbols, with about 6 bits per symbol, making a total of 8400 bits in the cell configuration. By way of comparison, Vogel [85], Bremermann [86], and others have suggested that the human genome contains 3×10^9 nucleotide pairs equivalent to 6×10^9 bits (at 2 bits per base pair), with roughly 500 base pairs per gene or 7×10^6 genes per haploid chromosome set. Bremermann points out that the human genome is roughly equivalent to 5000 books of 400 pages each, assuming the normal information content of English.

The chief difficulty in applying information theory to cell structure has been the lack of a mechanism, model, or theory which related the information content in bits of the code and the supposed information content of the cell. In the cell model one sees that there is no simple, and possibly no deterministic, relation-

ship between the bit content of a region of Turing enzyme automaton coding and the informational selective action this coding exerts in the cell, as measured by increasing the probabilities of certain states in a phase space or structures of the cell configuration. That is, the cell configuration pictured in figure 1 could in principle be used as an information coding device, with arbitrary assumptions of what aspects of the whole configuration would carry useful information. The problem would then be to relate the total enzyme automaton coding in bits to the information carrying capacity of the cell configuration in bits.

There would not seem to be any simple straightforward way of accomplishing this result, even though the cell is entirely algorithmic in operation, and even if it is not subjected to action of the environmental randomizing and leakage automata. Difficulties arise both from the standpoint of Turing machine coding and nature of the cell configuration. From experience with writing of hundreds of Turing automata, some of them with more than 1000 states, it has become perfectly evident to the author that Turing machine codes are just as arbitrary as any other computer codes and cannot be minimized in any unique way. That is, one can write a Turing machine to recode two adjacent binary 1's by a 0 with use of coding whose own information content might be a few dozen to many hundreds of bits, depending on programming strategy. In the case of enzyme automata the problem is much more complex.

A simpler standard test paradigm might be used. For example, all Turing cell automata might be required to function in a similar "style." Instead of having a cell configuration as complex as shown in figure 1, one might use string synthesizing automata to build a long chain polymer whose sidechains represented a binary signal, or special two dimensional patterns (as bar and square) standing for 1 and 0, or even use the positive and negative states of the genes in the cell model for information transmission. The point of this is to convert machine coding bits into configuration bits. However, numerous arbitrary assumptions would be required in such a highly artificial situation. A quite similar problem is encountered in transmission of words or numbers on TV, in which there is no obligatory relationship between the information content of the actual written message and the minimal numbers of bits needed, above a very low noisy limiting level given by Shannon's theory.

The cell is not a TV screen, but its genetic machinery and organelles define a level of organization greatly beyond the elementary thermodynamic one, and not necessarily well expressed by the apparent bit content of a sequence of DNA coding. The latter does represent a lower limit, as does channel capacity for a TV message transmission, but the actual level of organization is influenced by the coding of the cell configuration, entropy of the environment, mode of operation of enzyme automata and other factors. Lehninger [87] and Eigen and DeMaeyer [88] have analyzed information storage in general cellular structures, such as arrangements of lipids and proteins, in particular reference to memory.

The pertinent level of organization in this case is macromolecular, and not even molecular, to say nothing of atomic or electronic.

The applicability of quantum mechanical and statistical mechanics to cellular organization has been discussed fully by Elsasser [71], [72], who concludes that biological organization does not contradict any laws of quantum mechanics, but also is in no way predicted by quantum physics. The discussion above would suggest that one of the main problems is the organizational gap between the quantum level defined by an electron volt, the thermodynamic level defined by a Boltzman of energy, and the cellular level, which is perhaps defined by the hydrogen bond or ATP free energy level.

A quite similar dilemma arises in consideration of information processing in a computer regarded as a thermodynamic device. Even the pertinent energy levels needed to encode one bit, in presently available microcircuits or core memories, are many orders of magnitude removed from the thermodynamic bit level defined by the Boltzman equation

$$E_b = K_{Bz} \log_2 (p_2/p_1), \tag{10.1}$$

in which E_b is the energy of any elementary selection or informational representation process, shown by the probability ratio p_2/p_1, and K_{Bz} is the Boltzman constant, whose actual value is 1.38×10^{-16} ergs per degree. The Boltzman constant has the same dimensions as entropy. In the cell model it could be construed to mean the minimum unit of energy for either moving past one square of a finite grid containing figure 1, or moving one square on the cell contents tape traversed by the cellular enzyme Turing automaton. The cell model, TV screens, computers, and so forth, normally have much wasted motion compared to really impartant motion, and the limiting information packing density is never approached.

11. Conclusions

The cell model of this report bears somewhat the same relationship to real cells as the McCulloch-Pitts logical neuron has to a real nervous system: loose, nonquantitative, but based on fundamental mathematical concepts, and generally provocative. It provides a tentative framework for study of questions interrelating automata theory, information theory, molecular genetics and cell morphogenetics. It does not, and was not intended to, yield any quantitative facts about enzyme levels, and is at least four to five orders of magnitude simpler than even the simplest PPLO organism.

The type of self-organization in the cell model is quite different from those appropriate for neural network adaptive systems. The described measures of organization, namely, total time and energy for reproduction, are not appropriate for analysis of computers, memory networks, and previously described abstract self-organizing models. They would seem, however, to be very reason-

able measures of reproductive efficiency for real biological cells or tissues. An additional result achieved with the Turing enzyme automaton cell model is the partial clarification of certain troublesome questions pertaining to informational content of cell coding and apparent information stored in cell structures.

The presented model is best regarded as a new conceptual aid for understanding real cells and will no doubt be subject to much future modification and improvement. It does demonstrate that modern theoretical automata theory can be applied to molecular biology in a meaningful way.

REFERENCES

NOTE. Items shown with a JPRS number are Russian or German translations available from the U. S. Joint Publications Research Service, Office of Technical Services, U. S. Department of Commerce, Washington, D. C.

[1] W. R. STAHL and H. E. GOHEEN, "Molecular algorithms," *J. Theoret. Biol.*, Vol. 5 (1963), pp. 266–287.

[2] W. R. STAHL, "Self-reproducing automata," *Perspect. Biol. Med.*, Vol. 8 (1965), pp. 371–394.

[3] ———, "A model of self-reproduction based on string-processing finite automata," to appear in the *Proceedings of a Conference on Natural Automata and Useful Simulations*, sponsored by the Office of Naval Research, Stanford, June 1965.

[4] A. M. TURING, "On computable numbers, with an application to the Entscheidungsproblem," *Proc. London Math. Soc., Ser. 2*, Vol. 42 (1936), pp. 230–265.

[5] ———, "Correction to 'On computable numbers, with an application to the Entscheidungsproblem'," *Proc. London Math. Soc., Ser. 2*, Vol. 43 (1937), pp. 544–546.

[6] R. MCNAUGHTON, "The theory of automata," *Advances in Computers*, Vol. 2 (1961), pp. 379–421.

[7] B. A. TRAKHTENBROT, *Algorithms and the Machine Solution of Problems* (edited by S. V. Yablonskiy), Moscow, Gos. Izd. Fiz. Mat. Lit., 1960 (2nd ed.) (in Russian). (Available as *Algorithms and Automatic Computing Machines*, Boston, Heath, 1963.)

[8] M. A. ARBIB, *Brains, Machines, and Mathematics*, New York, McGraw-Hill, 1964.

[9] W. S. MCCULLOCH and W. PITTS, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, Vol. 5 (1943), pp. 115–133.

[10] R. W. COFFIN, H. E. GOHEEN, and W. R. STAHL, "Simulation of a Turing machine on a digital computer," *Proceedings of the AFIPS 1963 Fall Joint Computer Conference*, Baltimore, Spartan, 1963, pp. 35–43.

[11] W. R. STAHL, "Unsolvable problems for a cell automaton," *J. Theoret. Biol.*, Vol. 8 (1965), pp. 371–394.

[12] L. CHIRAVIGLIO, "Sequential machines and the coding of polypeptides," *J. Theoret. Biol.*, Vol. 8 (1965), pp. 130–140.

[13] B. CHANCE, J. H. HIGGINS, and D. GARFINKEL, "Analogue and digital computer representations of biochemical processes," *Fed. Proc.*, Vol. 21 (1962), pp. 75–86.

[14] D. GARFINKEL and B. HESS, "Metabolic control mechanisms VII: a detailed computer model of the glycolytic pathway in ascites cells," *J. Biol. Chem.*, Vol. 239 (1964), pp. 971–983.

[15] J. BONNER, *The Molecular Biology of Development*, Oxford, Oxford University Press, 1965.

[16] J. D. WATSON, *Molecular Biology of the Gene*, New York, Benjamin, 1965.

[17] H. J. MOROWITZ, *Life and the Physical Sciences; Introduction to Biophysics*, New York, Rinehart and Winston, 1963.

[18] F. W. STAHL, *The Mechanics of Inheritance*, Englewood Cliffs, Prentice-Hall, 1964.

[19] J. PAUL, *Cell Biology*, Stanford, Stanford University Press, 1964.
[20] F. JACOB and J. MONOD, "Elements of regulatory circuits in bacteria," *Biological Organization at the Cellular and Supercellular Level* (edited by R. J. C. Harris), New York, Academic Press, 1963, pp. 1–24.
[21] J. P. CHANGEUX, "The control of biochemical reactions," *Sci. Amer.*, Vol. 212 (1965), pp. 36–45.
[22] H. EAGLE, "Metabolic controls in cultured mammalian cells," *Science*, Vol. 148 (1965), pp. 42–51.
[23] H. J. MOROWITZ, W. A. HIGGINBOTHAM, S. W. MATTHYSSE, and H. QUASTLER, "Passive stability in a metabolic network (in cell reproduction)," *J. Theoret. Biol.*, Vol. 7 (1964), pp. 98–111.
[24] B. COMMONER, "Roles of deoxyribonucleic acid in inheritance," *Nature*, Vol. 202 (1964), pp. 960–968.
[25] E. ZUCKERKANDL and L. PAULING, "Molecules as documents of evolutionary history," *J. Theoret. Biol.*, Vol. 8 (1965), pp. 357–366.
[26] H. H. PATTEE, "The recognition of hereditary order in primitive chemical systems," *The Origins of Prebiological Systems and of Their Molecular Matrices*, New York, Academic Press, 1965, pp. 385–405.
[27] ———, "Molecular automata: a condition for the origin and evolution of life," to appear in the *Proceedings of a Conference on Natural Automata and Useful Simulations*, sponsored by the Office of Naval Research, Stanford, June 1965.
[28] J. D. BERNAL, "Molecular structure, biochemical function, and evolution," *Theoretical and Mathematical Biology* (edited by T. H. Waterman and H. J. Morowitz), New York, Blaisdell Publishing Co., 1965.
[29] H. J. MOROWITZ and M. E. TOURTELLOTTE, "The smallest living cells," *Sci. Amer.*, Vol. 206 (1962), pp. 117–126.
[30] H. J. MOROWITZ, "Cells: living; molecular physics," *Medical Physics*, Vol. 3 (edited by O. Glasser), Chicago, Year Book Publishers, 1960.
[31] G. E. HUTCHINSON, "The influence of the environment," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 51 (1964), pp. 930–934.
[32] J. VON NEUMANN, "The general and logical theory of automata," *Cerebral Mechanisms and Behavior* (edited by L. A. Jeffress), New York, Wiley, 1951, pp. 1–31.
[33] E. F. MOORE, "Mathematics in the biological sciences," *Sci. Amer.*, Vol. 211 (1964), pp. 148–164.
[34] A. W. BURKS, "Computation, behavior, and structure in fixed and growing automata," *Behavl. Sci.*, Vol. 6 (1961), pp. 5–22.
[35] J. MYHILL, "The possibilities of self-improvement in self-reproducing automata," notes from University of Michigan Summer Session on Automata Theory, 1963.
[36] M. A. ARBIB, "Some comments on self-reproducing automata," presented to the *Conference on Computer Science and Systems*, London, Ontario, September 1965, in press.
[37] ———, Personal communication regarding "Simple self-reproducing universal automaton," 1965.
[38] YA. M. BARZDIN', "Problems of universality in the theory of growing automata," *Dokl. Akad. Nauk SSSR*, Vol. 57 (1964), pp. 542–545 (in Russian).
[39] L. LOFGREN, "Kinematic and tesselation models of self-repair," *Biological Prototypes Synthetic Systems; Proceedings of the Second Annual Bionics Symposium, 1961* (edited by E. E. Bernard and M. R. Kare), New York, Plenum Press, 1962, Vol. 1, pp. 342–349.
[40] J. VON NEUMANN, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," *Automata Studies*, Princeton, Princeton University Press, 1956.
[41] S. WINOGRAD and J. D. COWAN, *Reliable Computation in the Presence of Noise*, Cambridge, Massachusetts Institute of Technology Press, 1963.
[42] J. D. COWAN, "Many-valued logics and reliable automata," *Principles of Self-Organization—1961*, New York, Pergamon Press, 1961.

[43] M. C. YOVITS and S. CAMERON (eds.), *Self-Organizing Systems—1959; International Tracts in Computer Science and Technology and Their Application*, Vol. 2, New York, Pergamon Press, 1960.

[44] M. C. YOVITS, G. T. JACOBI, and G. D. GOLDSTEIN (eds.), *Self-Organizing Systems—1962*, Washington, Spartan, 1962.

[45] E. E. BERNARD and M. R. KARE (eds.), *Biological Prototypes and Synthetic Systems; Proceedings of the Second Annual Bionics Symposium*, Vol. 1, New York, Plenum Press, 1962.

[46] H. VON FOERSTER and G. W. ZOPF, JR. (eds.), *Principles of Self-Organization—1962*, New York, Pergamon Press, 1962.

[47] V. M. GLUSHKOV, "Self-organizing systems and the abstract theory of computers," *Z. Vychisl. Mat. i Mat. Fiz. SSSR*, Vol. 2 (1962), pp. 459–466 (in Russian; JPRS: 14624).

[48] M. L. TSETLIN, "Finite automata and the modelling of simplest forms of behavior," *Uspehi Mat. Nauk*, Vol. 18 (1964), pp. 4–112.

[49] L. M. GEL'FAND, I. I. PYATETSKII-SHAPIRO, and M. L. TSETLIN, "Certain classes of games and games for automata," *Dokl. Akad. Nauk SSSR*, Vol. 152 (1963), pp. 845–848 (in Russian; JPRS: 22519).

[50] U. M. BORODYANSKY, "Some remarks on simulation of adaptation of environment of finite automata," *Kibernetika (Kiev)*, Vol. 1 (1965), pp. 29–34 (in Russian).

[51] A. A. LETICHEVSKIY and A. A. DORODNITSYNA, "Principles of the design of self-learning systems," Kiev, USSR, 1963 (in Russian; JPRS: 18181).

[52] W. R. ASHBY, "Principles of the self-organizing system," *Principles of Self-Organization—1962*, New York, Pergamon Press, 1962, pp. 255–278.

[53] W. R. ASHBY and J. RIGUET, "The avoidance of overwriting in self-organizing systems," *J. Theoret. Biol.*, Vol. 1 (1961), pp. 431–439.

[54] M. D. MESAROVIC, "On self-organizational systems," *Self-Organizing Systems—1962* (edited by M. C. Yovits, G. T. Jacobi, and G. D. Goldstein), Washington, Spartan, 1962, pp. 9–36.

[55] C. A. MUSES, "The logic of biosimulation," *Aspects of Theory of Artificial Intelligence* (edited by C. A. Muses and W. S. McCulloch), New York, Plenum Press, 1962, pp. 115–163.

[56] H. VON FOERSTER, "On self-organizing systems and their environments," *Self-Organizing Systems—1959* (edited by M. C. Yovits and S. Cameron), London, Pergamon Press, 1960, pp. 31–50.

[57] A. M. ANDREW, "The utilization of environment continuity by self-organizing systems," *Mathematical Theory of Automata Symposium* (edited by J. Fox), Brooklyn, Polytechnic Press, 1963, pp. 623–635.

[58] H. J. BREMERMANN, "Optimization through evolution and recombination," *Self-Organizing Systems—1962* (edited by M. C. Yovits, G. T. Jacobi, and G. D. Goldstein), Washington, Spartan, 1962, pp. 93–106.

[59] W. S. MCCULLOCH, "Biological computers," lecture at University of Pittsburg, March 1957.

[60] ———, "Stability of biological systems," *Homeostatic Mechanisms; Brookhaven Symposium Number 10*, New York, Brookhaven National Laboratory, 1958, pp. 207–216.

[61] ———, "Asymptotic nets combating noise," *Information Storage and Neural Control*, Springfield, Charles C Thomas, 1963, pp. 1–16.

[62] H. VON FOERSTER, "Molecular bionics," *Proceedings of a Bionics Symposium*, Wright-Patterson AFB, Aerospace Medical Division, Air Force Systems Command, ASD-TDR-63-946, 1964.

[63] S. BEER, *Cybernetics and Management*, New York, Wiley, 1964 (science edition).

[64] M. J. APTER and L. WOLPERT, "Cybernetics and development. I. Information theory," *J. Theoret. Biol.*, Vol. 8 (1965), pp. 244–257.

[65] G. M. FRANK, "Self-regulation of cellular processes," *Biological Aspects of Cybernetics—*

Collected Works, Moscow, Academy of Science, USSR, 1962 (in Russian; JPRS: 19637).

[66] G. DECHEV and A. MOSKONA, "Optimal self-regulation of metabolic processes in the cell; an attempt to establish a normal optimum of cell metabolism," *Dokl. Akad. Nauk SSSR*, Vol. 162 (1965), pp. 201–204 (in Russian).

[67] V. YA. ALEKSANDROV, "Excitation of living cells as a displacement in the stationary state of an open system," *Izv. Akad. Nauk SSSR, Ser. Biol.*, Vol. 3 (1963), pp. 621–629 (in Russian).

[68] K. A. KARIANI, "Some mechanisms of biochemical self-regulation at the cell level," *Biological Cybernetics—Collected Works*, Moscow, Academy of Science, 1962 (in Russian; JPRS: 19637).

[69] H. DAVID, "The cell as a cybernetic system," *Das Deutsche Gesundheitswesen*, Vol. 20 (1965), pp. 53–62; pp. 101–107 (in German; JPRS: 28974).

[70] E. SCHRODINGER, *What Is Life?*, New York, Macmillan, 1945.

[71] W. M. ELSASSER, *The Physical Foundation of Biology*, London, Pergamon Press, 1958.

[72] ———, "Synopsis of organismic theory," *J. Theoret. Biol.*, Vol. 7 (1964), pp. 53–67.

[73] R. ROSEN, "A note on the quantum-theoretic basis of primary genetic activity," *Bull. Math. Biophys.*, Vol. 25 (1963), pp. 183–187.

[74] H. C. LONGUET-HIGGINS, "Quantum mechanics and biology," *Biophys. J.*, Vol. 2 (1962), pp. 207–215.

[75] G. KARREMAN, "Towards quantum biology," *Biopolymers Symposium* (edited by M. Weissblüth), New York, Wiley, 1964, Vol. 1, pp. 313–324.

[76] B. C. GOODWIN, *Temporal Organization in Cells*, New York, Academic Press, 1963.

[77] K. S. TRINCHER, "Thermodynamic theorem of biological evolution," *Biofizika*, Vol. 7 (1962), pp. 740–744.

[78] ———, "On modeling the functions of a living system," *Cybernetics, Thought, and Life* (edited by A. K. Berg *et al.*), Moscow, 1964, pp. 206–220 (in Russian).

[79] A. D. URSUL, "On the nature of information," *Voprosy Filosofii*, Vol. 3 (1965), pp. 131–140 (in Russian).

[80] W. R. STAHL, R. W. COFFIN, and H. E. GOHEEN, "Simulation of biological cells by systems composed of string-processing finite automata," *Proceedings of the AFIPS 1964 Spring Joint Computer Conference*, Baltimore, Spartan, 1964, pp. 89–102.

[81] A. L. SKOULTCHI and H. J. MOROWITZ, "Information storage and survival of biological systems at temperatures near absolute zero," *Yale J. Biol. Med.*, Vol. 37 (1964), pp. 158–163.

[82] H. P. YOCKEY, R. L. PLATZMAN, and H. QUASTLER (eds.), *Symposium on Information Theory in Biology*, New York, Pergamon Press, 1958.

[83] L. L. GATLIN, "The information content of DNA," typescript.

[84] H. A. JOHNSON, "Redundancy and biological aging," *Science*, Vol. 141 (1963), pp. 910–912.

[85] F. VOGEL, "A preliminary estimate of the number of human genes," *Nature*, Vol. 201 (1964), p. 847.

[86] H. BREMERMANN, "Analysis of self-organizing systems," to appear in *Proceedings of the Ninth Biophysical Society Meeting*. (See also "Quantitative aspects of goal-seeking, self-organizing systems," *Progress in Theoretical Biology*, Vol. 1, New York, Academic Press, 1967.

[87] A. L. LEHNINGER, "Supramolecular organization of enzyme and membrane systems," *Neurosciences Research Bulletin*, Vol. 3 (1965), pp. 26–47.

[88] M. EIGEN and L. C. M. DEMAEYER, "Information and storage in biomolecular systems," *Neurosciences Research Bulletin*, Vol. 3 (1965), pp. 1–24.