

## **System Theoretic Analysis of Models: Computer Simulation of a Living Cell**

BERNARD P. ZEIGLER AND ROGER WEINBERG

*Department of Computer and Communication Sciences,  
Logic of Computers Group, The University of Michigan,  
611 Church Street, Ann Arbor, Mich. 48104, U.S.A.*

*(Received 2 December 1969, and in revised form 23 January 1970)*

In modeling a biological system for computer simulation one must often reduce the complexity of the system in accord with practical considerations of computer size and length of computation time. In our model of a living cell (Weinberg & Zeigler, 1969) we found the formal concepts of systems theory valuable in this regard. In particular, the notion of homomorphism is applicable to the problem of assuring that the model preserves interesting properties of the system under study while at the same time being simpler with respect to computer implementation.

To develop our model of the cell we simplified the state space and transition function by aggregating co-ordinates of the state description while at the same time reducing the state sets of the new lumped co-ordinates. Certain aspects of the transition function of the model could be determined beforehand, but full specification of the transition function had to be determined by trial and error. In general, testing the model for homomorphism is equivalent to validating the model by comparing its behavior with the behavior of its real world counterpart.

The process of aggregation by which we simplified our model, while well-known to economists, has been little appreciated by simulators of biological systems. This paper discusses the application of these system theory concepts to biological models for computer simulation.

### **1. Introduction**

The problem immediately confronting any attempt to model a biological system for computer simulation is the enormous complexity inherent in such systems. In this paper we present an approach to this problem based upon the concepts of homomorphism and aggregation. A mathematical formulation is developed to discuss the application of these concepts to a model of a living *Escherichia coli* cell which has been simulated on a digital computer and successfully tested against laboratory data (Weinberg, 1968*a,b*; Weinberg & Berkus, 1969*a,b*; Weinberg & Zeigler, 1970).

Klir & Valach (1967) discuss the important connection between simulation and model construction and the mathematical notion of homomorphism. One may attempt to achieve such homomorphic mappings by lumping entities together—this constitutes the basic character of our model. The importance of this process of aggregation has been known to economists for some time (e.g. Green, 1965; Ijiri, 1968). However, the concepts of aggregation and homomorphism have not been much appreciated by simulators and modellers, especially those of biological systems. [They are not mentioned, for example, in the extensive discussion of Fein (1966), nor by Rosen (1968).] In this paper we show that at the conceptual level, these concepts help explicate the notions of model complexity and adequacy. At the more practical level, they suggest and justify techniques for the reduction of complexity and for the testing of model adequacy in the context of biological system simulations.

Simon & Ando (1961) presented a mathematical formulation of aggregation of variables for dynamic systems and studied a particular class of systems (nearly decomposable systems) whose long term behavior satisfied these conditions. An informal discussion of the relationship of these systems to the complexity of natural systems was given by Simon (1962).

Since 1961, there has arisen an extensive study of the structural and behavioral relations which hold between automata (Hartmanis & Stearns, 1968; Krohn & Rhodes, 1965; Arbib, 1969; Zeigler, 1968) and the relation of automata to general systems (Kalman, Falb & Arbib, 1968). With these developments it is possible to give a unified mathematical formulation of aggregation and its relation to the more basic notion of the homomorphism of systems. Based on this formulation and work in system complexity (Zeigler, 1968, 1969) one can demonstrate precisely how the complexity (as determined by measures relevant to computer implementation) can be reduced in going from a system to a homomorphic image. In this paper we give only the basics of this formulation necessary to analyze the basic character of our living cell model. A more complete formal presentation is planned (Zeigler, manuscript in preparation). Hopefully, this analysis will motivate and justify the application of these concepts to other biological models and simulations.

## 2. Model of *E. coli* Cell

### (A) BASIC MODEL CHARACTERISTICS

In writing a simulation of the *E. coli* cell (Weinberg 1968*a,b*) we recognized that in even the simplest of cells there are more than 3000 different kinds of molecules in an intricate spatial and functional relationship. These molecules co-exist in a complex metabolic network with superposed genetic and enzy-

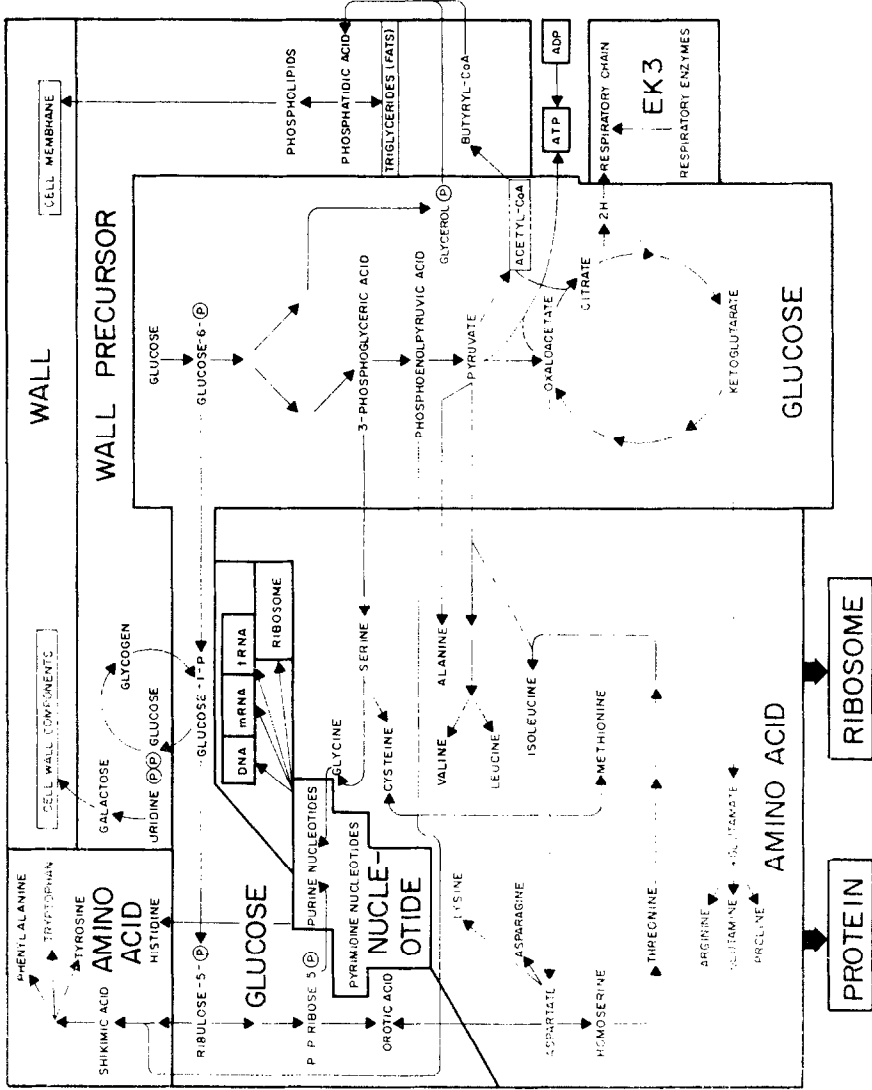


Fig. 1. Some of the primary molecular constituents and metabolic pathways in the *E. coli* cell are shown. These constituents are shown grouped together into "pools" e.g., GLUCOSE, AMINO ACIDS, etc., which serve as the entities for a model of the cell.

matic control. This complexity had to be reduced drastically to enable a model to be simulated on a digital computer. It is important to note that while electronic computers continue to attain increasingly greater memory stores and rates of operation, this capacity is very much limited when one considers simulating *in detail* the operation of a natural system.

To construct our model we created the metabolic network consisting of a small number of chemical pools as shown in Fig. 1. The pools were to represent the network behavior in an aggregate manner. As shown in Fig. 2, the metabolic and informational transactions established among the pools was to reflect in a coarser way the underlying relations among their constituents. [A detailed description of the basic model is available (Weinberg & Berkus, 1969*a,b*). Some results of simulation are reported by Weinberg & Zeigler (1969).]

This represents a novel approach to models of the metabolism of the cell since most simulation studies (Garfinkel, 1966; Yeisley & Pollard, 1964; Savageau, 1969) are concerned with simulating particular pathways in detail rather than attempting to simulate the global cellular operation. Heinmets (1966) has written a model system for a generalized cell which attempts a global analysis. However, the model was not intended to enable comparison of model and real cell behavior. The possibility of such a comparison is an essential objective of our simulation.

Certain considerations guided our assignment of molecules into pools:

*Metabolic topology* led us to combine into a single entity only those molecules which could be drawn adjacent to one another on the metabolic

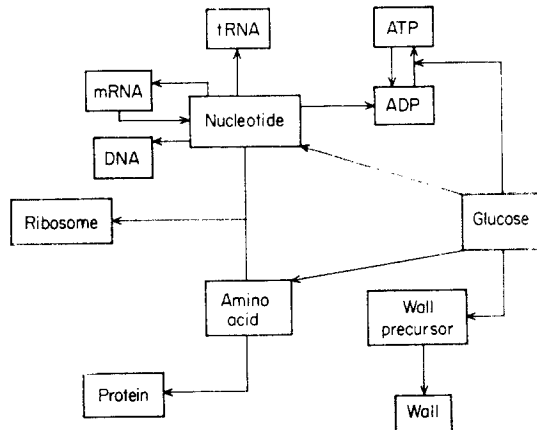


FIG. 2. The primary chemical pools used in the model are shown. The arrows indicate the direction of the flow of material between pools. The topology of this flow preserves the metabolic topology of Fig. 1, in a manner explained in the text.

map (Figs 1 and 2). For example, in a pathway  $A \rightarrow B \rightarrow C$ , chemicals A, B, C might be partitioned into the pools A and BC, but not into the pools AC and B.

*Functional relationships* between groups of molecules were extremely important, and the molecules lumped together in any one model entity constituted, in some way, a functional unit (Table 1). Thus, all the molecules produced in the exothermic breakdown of sugar to carbon dioxide and water were lumped in this model since they could be considered functionally as molecules intermediate in a chemical pathway used to produce energy. Later models may employ more refined partitions, for example to capture the subtle and important relationships between molecules at different points in the glycolytic pathway, citric acid cycle and cytochrome system.

TABLE 1

*Possible state spaces and transition functions for models of the living cell*

State space	Co-ordinates	Co-ordinate sets	Transition function
Atonic	Electrons, nuclei	Position, momentum	Schrodinger's equation (quantum mechanics) e.g.
Molecular	Molecules	Shape, active site, energy level	Koch (1967) probabilistic model
Concentration	Molecule type e.g. ADP, DNA	Number of molecules of ADP/cell etc.	Mathematical and logical equations based on chemical kinetics e.g. Chance <i>et al.</i> (1965)
Higher level groupings	Biochemical pools e.g. amino acids	number of molecules of pool/cell	Mathematics and logical equations based on chemical kinetics and Molecular control mechanisms e.g. repression, allosteric inhibition (Weinberg, 1969; Goodwin, 1969)

The co-ordinates at one level are aggregations of the co-ordinates at a lower level of organization.

*Experimental data* were often available for large chemical pools, e.g. products of glycolysis and the citric acid cycle, making these separate entities logical candidates for grouping into single entities in this model of the cell.

It is clear that, as intended, our model is a representation of a real cell from which much has been omitted. Indeed, any model amenable to human and/or computer analysis must neglect many more features than it considers, and in such abstractions we run the risk of losing any meaningful relation to the "real thing".

Our model of the cell differs from others in that the abstraction involved is that arising from aggregation of variables rather than selection of subsystems. In order to discuss the general application of such a technique to biological systems and its ability to produce valid models we now introduce some elementary system theoretic concepts.

#### (B) SYSTEM THEORETIC FORMULATION OF MODEL

In its most basic form, a *system*<sup>†</sup> is defined as a set of states  $S$ , together with a *transition function*  $\tau : S \rightarrow S$ .  $\tau$  describes the behavior of the system over time by indicating which next state is to follow the present state. Thus if the state at time  $t$  is  $s(t)$  then the state at time  $t + 1$ ,  $s(t+1) = \tau[s(t)]$  where  $t = 0, 1, 2, \dots$

The *state space*  $S$  is usually described as a cartesian product of component state sets i.e.  $S = \prod_{\alpha \in D} S_{\alpha}$  where  $D$  is a finite set of *co-ordinates* (or entities) and  $S_{\alpha}$  is the state set (or attribute set) of co-ordinate  $\alpha$ . In Table 1 we list a number of possible state spaces and indicate the form a transition function might take in each. In Table 2 and Fig. 3 we specify in more detail the state space and transition function of the present model. Thus, the co-ordinates of the model include the chemical pools of Fig. 2 (amino acids, protein, glucose, etc.), the enzymes ( $Ek_1, Ek_2, \dots, Ek_{10}$ ), and the different messenger RNA's ( $RNk_1, \dots, RNk_{10}$ ). We must also keep track of volume of the cell, the number of cells, and information to be used by the replication subroutine. Accordingly, these are also represented in our state vector.

The transition function for our model consists of difference equations and Boolean expressions describing: (1) enzyme catalyzed chemical reactions, (2) allosteric modification of enzymes, (3) repression of messenger RNA

<sup>†</sup> Our definition of "system" essentially is that routinely employed in automata theoretic formulations. (For example, Hartmanis & Stearns, 1968.) A more general definition (incorporating continuous time) is given by Kalman, Falb & Arbib (1969, pp. 5 to 12). Arbib (1969, pp. 51 to 57) traces the steps involved in restriction to discrete time systems and finite state systems.

production, (4) self-replication of DNA under genetic controls, and (5) permeability of the cell to the chemical pools represented in the simulation (Table 2). The transition function captures relationships not apparent in the state co-ordinate representation (Fig. 3).

Noting that  $S = \times_{\alpha \in D} S_\alpha$ ,  $\tau$  is equivalent to a set of maps  $\{\tau_\beta | \beta \in D\}$  where each  $\tau_\beta$  maps  $S$  to  $S_\beta$ . The co-ordinate function  $\tau_\beta$  tells how the next state of co-ordinate  $\beta$  depends on the present state of the system. (Clearly, if we tell

TABLE 2  
*Co-ordinates and co-ordinate sets for model of living cell*

Co-ordinate = Entity	Range of value of co-ordinates Attribute of entity
Pools of chemicals, PRDC(1), ..., PRDC(10)	Concentration of pool
Enzymes, EK(1), ..., EK(10)	Concentration of enzyme
Messenger RNA, RNK(1), ..., RNK(10)	Concentration of RNA
Genetic apparatus	Amount of DNA, site of replication, number of genes in cell for producing sites for replication
Cell volume	Total volume of the cell
Cell number	Number of cells represented in the culture

The function for calculating the state of system in the next time step from the present state is outlined as follows:

A. *The differential and Boolean equations* relating concentrations of variables at a given time to the concentrations of those variables DT seconds later. E.g. for AA, the amino acid pool, one needs enzyme EK(2) to catalyze the production of AA from glucose, and one uses ATP as an energy source. At the same time, AA is lost as it is used for the production of RIB and PRTN.

1.  $DAA = K(2)*GLUC*DT*EK(2)*ATP - 1.E6/102.*DRIB - (4.E4/102.)*DPRTN$   
     production of AA from                      loss of AA to                      loss of AA to PRTN  
     GLUC    RIB
2. RNK(2) produced EK(2) from AA under the direction of DNA, using ATP for energy
3.  $DEK(2) = K(7)*AA(RNK(2)/MRNAO)*DT*EK(7)*ATP$ .
4. RNK(2) itself was produced from NUC under the direction of DNA, catalyzed by EK8, using ATP for energy. RNK(2) decayed spontaneously at the same time, producing some loss of RNK(2) already present.
5.  $DRNK(2) = (K8K(2)*NUC*DNA*EK(8)*ATP - KDRNK*RNK(2))*DT$   
     production of RNK(2)    decay of RNK(2)

- B. *Allosteric modification* of enzymes simulated by modifying the rate constant which characterizes all different forms of any enzyme associated with a particular reaction.
- C. *Repression* of messenger RNA directing the production of a particular enzyme.
- D. *Genetic behavior* of DNA in response to the state of the cell.
- E. *Permeability*.

how each co-ordinate is to change state we will be specifying how the system changes state, and conversely). Indeed,  $\tau_\beta$  may not depend on all the co-ordinates in  $D$ : let  $I_\beta$  be the subset of  $D$  on which  $\tau_\beta$  depends, then  $\tau_\beta$  maps  $\times_{\alpha \in I_\beta} S_\alpha$  into  $S_\beta$ .†

Based on the sets  $I_\beta$  we can draw a *digraph* (directed graph) which represents the coordinate interdependence in the system. The digraph, which we also denote by  $D$ , has as points the set  $D$  and we draw a line from  $\alpha$  to  $\beta$  just in case  $\alpha \in I_\beta$ .

As an example, consider the schematic view of some of the main metabolic pathways in *E. coli* of Fig. 1. Under conditions where the enzyme concentration remains fixed‡ we can obtain the digraph,  $D^E$  of the system with concentration state space  $S^E$  directly as follows:  $D^E$  is the set of chemicals shown in Fig. 1. For each of the lines shown we must add a line in the opposite direction; further, when any two points are adjacent to a third they must be joined by lines in both directions. (These rules are inferred from the chemical kinetic equations.) For example, the input set  $I_{\text{Gluc}}$ , of glucose-1-P consists of glucose, glucose-1-P and glycogen. This means that  $\tau_{\text{Gluc}}$  which specifies the next concentration of this chemical depends only on the concentrations of the chemicals in  $I_{\text{Gluc}}$ .

We mention that the interrelation between properties of the digraph (structure) and properties of the transition function (behavior) have been fruitfully applied to biochemical systems by Kauffman (1969) and others.§

† Given any subset  $T \subseteq \times_{\alpha \in D} S_\alpha$ , a co-ordinate  $\alpha$  depends on a set  $I_\alpha^T$  if  $(\forall s, s' \in T) (\forall \beta \in I^T)$  [proj $_\beta(s) = \text{proj}_\beta(s') \Rightarrow \tau_\alpha(s) = \tau_\alpha(s')$ ] and no proper subset of  $I_\alpha^T$  has this property. This definition is that given in Zeigler (1968, 1969) and the well-known partition pair calculus of Hartmanis & Stearns (1966). By way of explanation, this criterion selects a minimal set  $I_\alpha^T$  such that a function defined only on  $\times_{\beta \in I^T} S_\beta$  can be found which equals  $\tau_\alpha$  on  $T$ .

Assuming a finite number of co-ordinates (though not necessarily a finite state set) will guarantee the existence of such sets; more generally a minimum condition is needed.

We are not guaranteed however that for each  $\alpha$  there is only one set  $I_\alpha^T$  on which it depends. Nevertheless, when  $T = \times_{\alpha \in D} S_\alpha$  these sets are unique. This situation is discussed by Zeigler (1968) and Roosen Runge (1967).

For each choice of sets  $\{I_\alpha^T | \alpha \in D\}$  we can construct an associated digraph  $D^T$ . Thus to each subset of the state space there corresponds one (or more) well-defined digraphs. In particular, there is a unique well-defined digraph associated with the whole space  $\times_{\alpha \in D} S_\alpha$ .

‡ Of course the situation is much more complicated than this and our complete model will represent this fact. For example, co-ordinate interdependencies arising from the allosteric inhibition and repression components of the transition function are reflected in the digraph.

§ Kauffman's state space is the discrete set of all genome off-on configurations. This is an abstraction of the state space of our model. Our discussion of homomorphism can be used to relate the two models.



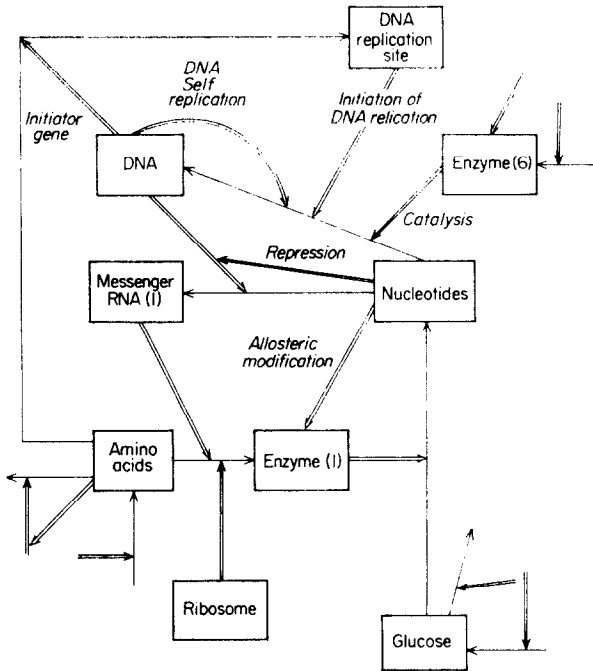


FIG. 3. Some of the relationships which are embodied in the equations of the transition function of the cell model are shown. These relationships, repression, allosteric modification, etc., are shown by a double arrow ( $\Rightarrow$ ) to distinguish them from the flow of materials indicated by a single arrow ( $\rightarrow$ ).

### 3. Homomorphism and Model Simplification

#### (A) HOMOMORPHISM OF SYSTEMS

The notion of homomorphism provides a criterion by which the validity of a model may be judged.

Let  $S, S'$  be state spaces and  $\tau, \tau'$  their associated transition functions. We seek conditions under which the system  $(S', \tau')$  can be reasonably said to be a model or simplification of  $(S, \tau)$ .

To do this, we say that a *trajectory* through  $S$  beginning at  $s$  in  $S$  is any sequence

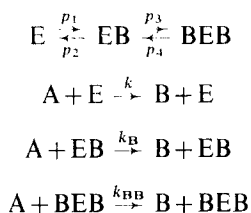
$$\langle s, \tau(s), \tau^2(s), \dots, \tau^n(s) \rangle, \quad \text{where } \tau^i(s) = \tau^{i-1}[\tau(s)], \tau^1(s) = \tau(s).$$

The set of such trajectories is the *behavior* of the system  $(S, \tau)$ . Let  $h : S \rightarrow S'$  be a map from  $S$  onto  $S'$ ; we want to preserve the behavior of  $(S, \tau)$  in the sense that for any trajectory  $\langle s, \tau(s), \tau^2(s), \dots, \tau^n(s) \rangle$  in the behavior of  $(S, \tau)$ ,  $\langle h(s), h[\tau(s)], h[\tau^2(s)], \dots, h[\tau^n(s)] \rangle$  is a trajectory in the behavior of  $(S', \tau')$ .

This is equivalent to the condition that for every  $s \in S$ ,  $h[\tau(s)] = \tau'[h(s)]$ . In practice this condition may be required to hold only approximately and then only for that part of the state space which may be explored experimentally. If  $h$  satisfies this condition it is said to be a *homomorphism* and  $(S', \tau')$  is a homomorphic image of  $(S, \tau)$ . When  $h$  is a one-one homomorphism, it is an *isomorphism* and  $(S', \tau')$  is isomorphic to  $(S, \tau)$ . We shall take as the formal counterpart of the statement that “ $(S', \tau')$  is a model of  $(S, \tau)$ ” the statement “ $(S', \tau')$  is a homomorphic image of  $(S, \tau)$ ” our justification being that the notion of “model” entails the concept of behavior preservation we have given.

(B) APPLICATION TO ALLOSTERIC INHIBITION

We shall illustrate the homomorphism concept with a technique actually employed in our model. Consider a subsystem consisting of substrate A, product B, and enzyme E which can exist in three forms: E (pure enzyme), EB (one-molecule of B attached) and BEB (two molecules of B attached). The chemical kinetics of this system are given by:



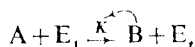
where the  $p$ 's and  $k$ 's are suitable rate constants. Call this system  $S$ . Then  $S$  has as co-ordinates  $D = \{A, B, E, EB, BEB\}$  with state space  $S = (A) \times (B) \times (E) \times (EB) \times (BEB)$  where for example,  $A$  denotes the concentration of A and  $(A)$  is the range over which  $A$  varies. The transition function  $\tau$  is specified by the co-ordinate functions  $\{\tau_A, \tau_B, \tau_E, \tau_{EB}, \tau_{BEB}\}$  namely,

$$\begin{aligned} \tau_A(A, E, EB, BEB) &= A - k_A E A - k_B E B \cdot A + k_{BEB} BEB \cdot A \\ \tau_B(A, B, E, EB, BEB) &= B + k_A E A + k_B E B \cdot A + k_{BEB} BEB \cdot A \\ \tau_E(B, E, EB) &= E - p_1 E \cdot B + p_2 EB \quad (1) \\ \tau_{EB}(B, E, EB, BEB) &= EB + p_1 E \cdot B - p_2 EB - p_3 EB \cdot B + p_4 BEB \\ \tau_{BEB}(B, EB, BEB) &= BEB + p_3 EB \cdot B - p_4 BEB. \end{aligned}$$

(These are discrete versions of the usual kinetic differential equations.)

We wish to simplify  $S$ , i.e. to replace it with a homomorphic image system  $S'$ . Let  $S'$  have as co-ordinates  $\{A, B, E_i\}$  where  $E_i$  is to represent the aggregate of the enzyme in its three forms. Then  $S'$  has as state space  $S' = (A) \times (B) \times (E_i)$  and we define a map  $h: S \rightarrow S'$  given by  $h(A, B, E, EB, BEB) = (A, B, E + EB + BEB)$ .

The problem now is to define a transition function  $\tau'$  for  $S'$  which will make  $h$  a homomorphism. To do this we shall consider the kinetics of  $S'$  to be given by:



where  $K(B)$  will be a rate "constant" depending on  $B$  thus reflecting the feedback effect of allosteric inhibition. The transition equations of system  $S'$  then are:

$$\begin{aligned} \tau'_A(A, E_i) &= A - K(B) \cdot E_i \cdot A \\ \tau'_B(A, B, E_i) &= B + K(B) \cdot E_i \cdot A \\ \tau'_{E_i}(E_i) &= E_i. \end{aligned} \tag{2}$$

The condition under which  $h$  is a homomorphism is then that

$$h[\tau(A, B, E, EB, BEB)] = \tau'[h(A, B, E, EB, BEB)].$$

By working out the A co-ordinate we obtain the equation:

$$A(1 - kE - k_B EB - k_{BB} BEB) = A[1 - K(E + EB + BEB)].$$

A similar equation results in the B position. An equation involving  $E, EB, BEB$  appears in the  $E_i$  position which is trivially satisfied (it is a result of the fact that in system  $S$  the total enzyme concentration  $E + EB + BEB$  is constant). Thus  $h$  will be a homomorphism just in case there is a function  $K(B)$  such that

$$kE + k_B EB + k_{BB} BEB = K(B) \cdot (E + EB + BEB) \tag{3}$$

for all values of *interest* of  $E, EB, BEB$ . [Clearly unless

$$K(B) = k = k_B = k_{BB},$$

which is not generally the case, the equation cannot hold for *all* values of  $E, EB, BEB$ .]

We shall shortly examine the factors determining the extent to which equation (3) can be satisfied.

Now under steady state conditions we assume that the concentrations of the enzyme forms  $E, EB, BEB$  of system  $S$  remain constant. This results in the following equations:

$$\begin{aligned} -p_1 \tilde{E} \cdot \tilde{B} + p_2 \cdot \tilde{EB} &= 0 \\ p_1 \tilde{E} \cdot \tilde{B} - p_2 \cdot \tilde{EB} - p_3 \cdot \tilde{BE} \cdot \tilde{B} + p_4 \cdot \tilde{BEB} &= 0 \\ p_3 \cdot \tilde{EB} \cdot \tilde{B} - p_4 \cdot \tilde{BEB} &= 0 \end{aligned} \tag{4}$$

where, for example  $\tilde{E}$  is the constant value of concentration  $E$  in a steady-state condition.

Manipulating the equations for equilibrium conditions, one can obtain the following expressions for  $\tilde{E}$ ,  $\widetilde{EB}$  and  $\widetilde{BEB}$ :

$$\begin{aligned}\tilde{E} &= \tilde{E}_{\text{total}} \cdot \frac{1}{(1 + P_1 \cdot \tilde{B} + P_3 \cdot \tilde{B}^2)} \\ \widetilde{EB} &= P_1 \cdot \tilde{B} \cdot \tilde{E} \\ \widetilde{BEB} &= P_3 \cdot \tilde{B}^2 \cdot \tilde{E}\end{aligned}\quad (5)$$

where

$$\begin{aligned}P_1 &= p_1/p_2 \\ P_3 &= p_3/p_4\end{aligned}$$

(To obtain these expressions, note that:

$$\begin{aligned}\widetilde{EB} &= (p_1/p_2) \cdot \tilde{B} \cdot \tilde{E} = P_1 \cdot \tilde{B} \cdot \tilde{E} \\ \widetilde{BEB} &= (p_3/p_4) \cdot \tilde{B}^2 \cdot \tilde{E} = P_3 \cdot \tilde{B}^2 \cdot \tilde{E} \\ \tilde{E} &= \tilde{E}_{\text{total}} - \widetilde{EB} - \widetilde{BEB} = \tilde{E}_{\text{total}} - \frac{p_1}{p_2} \cdot \tilde{B} \cdot \tilde{E} - \frac{p_3}{p_4} \cdot \tilde{B}^2 \cdot \tilde{E}.\end{aligned}$$

Dividing out  $\tilde{E}$  from both sides, one obtains

$$1 = \frac{-p_1}{p_2} \cdot \tilde{B} - \frac{p_3}{p_4} \cdot \tilde{B}^2 + \tilde{E}_{\text{total}}/\tilde{E}.$$

Manipulating, and substituting  $P_1 = p_1/p_2$  and  $P_3 = p_3/p_4$  one obtains

$$\tilde{E} = \tilde{E}_{\text{total}} \cdot \frac{1}{1 + P_1 \cdot \tilde{B} + P_3 \cdot \tilde{B}^2}.$$

Substituting these expressions in the homomorphism equation (3), cancelling and collecting terms yields:

$$k + (P_1 \cdot k_B) \tilde{B} + (P_3 \cdot k_{BB}) \tilde{B}^2 = K(\tilde{B})(1 + P_1 \cdot \tilde{B} + P_3 \cdot \tilde{B}^2). \quad (6)$$

We shall now show how a value of  $K(\tilde{B})$  can be estimated from experimental data for a given steady-state concentration value  $\tilde{B}$ . Steady-state concentrations of the pools are assumed to be known for given nutritional environments.

Rate constants for flow of materials between pools are calculated for each environment from the steady-state concentrations of the pools in that environment, the initial volume of the cell at the beginning of a generation, and the time necessary for all quantities in the cell to double in that environment, i.e. the time for one cell generation. Exponential increase in cell mass and in all pool amounts, except for a linear increase in DNA and genetic apparatus, is assumed. The calculations will be illustrated for protein production from amino acids given by:

$$\text{Protein} \xrightarrow{K} \text{Amino acid}$$

The chemical kinetic equation is:

$$\frac{d[\text{Protein}(t)]}{dt} = K[\text{Amino}(t)] \cdot \text{Volume}(t)$$

where

Protein ( $t$ ) = number of protein molecules in the cell at time  $t$ . =

Amino ( $t$ ) = the number of amino acid molecules per cell.

Volume ( $t$ ) = volume of cell at time  $t$  (set at 1 for cell at beginning of generation).

Integrating over one cell cycle, during which volume doubles

$$\int_{\text{Protein}(0)}^{\text{Protein}(1)} d[\text{Protein}(t)] = \int_{t=0}^{t=1} K[\text{Amino}(T)] \text{Volume}(T) dt$$

where

Protein (0) = total number of protein molecules in the cell at time = 0, the beginning of a cell cycle.

Protein (1) = total number of protein molecules in the cell at time = 1, the end of a cell cycle.

The increase in volume in a steady-state growth condition in one environment is just sufficient to keep pool concentrations constant, while increase in volume occurs exponentially at a rate permitting volume to double over one cell generation. The amino acid concentration in the cell, therefore remains constant in a given environment.  $K$  is also constant in this environment so that  $K$  and  $[\text{Amino}(t)]$  can be placed in front of the integration sign to give

$$\int_{\text{Protein}(0)}^{\text{Protein}(1)} d[\text{Protein}(t)] = K \cdot [\text{Amino}(t)] \cdot \int_{t=0}^{t=1} \text{Volume}(t) dt.$$

Since volume is assumed to increase exponentially during steady-state growth conditions, and  $\text{Volume}(0) = 1$ , one can set up the equation

$$\text{Volume}(t) = \text{Volume}(0) \cdot e^{at} = 1 \cdot e^{at} = e^{at}.$$

Since the volume of a cell doubles in one generation,  $\text{Volume}(1) = 2$ . Thus  $a = \ln \text{Volume}(1) = \ln 2$ . Going back to the integral relating increase in number of protein molecules to volume and amino acid concentration, and substituting  $e^{\ln(2) \cdot t}$  for  $\text{Volume}(t)$ , one obtains

$$\int_{\text{Protein}(0)}^{\text{Protein}(1)} d[\text{Protein}(t)] = K \cdot [\text{Amino}(t)] \cdot \int_{t=0}^{t=1} e^{\ln(2) \cdot t} dt.$$

Integrating, one obtains  $\text{Protein}(1) - \text{Protein}(0) = K \cdot [\text{Amino}(t)] / \ln(2)$ . All amounts double in one-cell cycle, therefore  $\text{Protein}(1) = 2 \cdot \text{Protein}(0)$ . The concentration of amino acid is equal to total amount of the amino acid pool per cell volume, i.e.  $[\text{Amino}(t)] = \text{Amino}(t) / \text{Volume}(t)$ . At  $t = 0$ ,

$[Amino(0)] = Amino(0)/Volume(0)$ . Since  $[Amino(t)]$  remains constant in an environment for steady-state growth,  $[Amino(t)] = Amino(0)/Volume(0)$ . With  $Volume(0) = 1$ ,  $[Amino(t)] = Amino(0)$  and one obtains  $Protein(0) = K \cdot Amino(0)/\ln(2)$ . Solving for  $K$ , one obtains  $K = [Protein(0)/Amino(0)] \cdot \ln(2)$ .

The amounts of the protein and amino acid pools at the beginning of a cell cycle,  $Protein(0)$  and  $Amino(0)$ , can be calculated from the data available in the literature on the average amounts of these pools. We will show that since the pools size increases exponentially during growth, the pool size at the beginning of a generation equals the average pool size times the natural logarithm of 2. The equation for calculating the average volume over one cell cycle is

$$\int_{t=0}^{t=1} Volume(t) \cdot Probability(t) dt = \text{average volume over a cell cycle}$$

where

- $t =$  time in life cycle,
- $t = 0$  at beginning of cell cycle,
- $t = 1$  at end of cell cycle,

$Volume(t) =$  volume of cell at time  $t$ ,

$Probability(t) =$  probability density for the time in the life cycle of a cell being considered.

We shall set  $Probability(t) = 1$  since bacterial growth is ergodic, and unsynchronized cultures were used for experimental data from the literature, giving the uniform density. Thus

$$\int_{t=0}^{t=1} Volume(0) \cdot e^{\ln(2) \cdot t} \cdot 1 dt = \text{average volume over a cell cycle.}$$

Carrying out the integration, average volume =  $Volume(0)/\ln(2)$ , and  $Volume(0) = \text{average volume} \cdot \ln(2)$ .

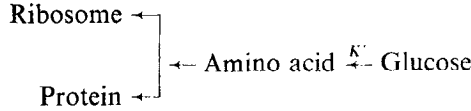
Similar equations indicate that the average pool size for a cell must be multiplied by  $\ln(2)$  to give the pool size at the beginning of a cell cycle. The equations for the amino acid pool are developed as an illustration.

$$\int_{t=0}^{t=1} Amino(t) \cdot Probability(t) dt = \text{average amino acid pool size}$$

In a steady-state environment  $Amino(t) = [Amino(0)] \cdot Volume(t)$ ,  $Volume(t) = e^{\ln(2)t}$  and  $Probability(t) = 1$ . Substituting for  $Amino(t)$  and  $Probability(t)$ , one obtains

$$\int_{t=0}^{t=1} Amino(0) \cdot e^{\ln(2) \cdot t} dt = \text{average amino acid pool size}$$

Integrating, one obtains average amino acid pool size = Amino (0) · ln (2).  
 From the assumed metabolic pathway:



one can set up the equation

$$\frac{d[\text{Amino}(t)]}{dt} = K' \cdot [\text{Glucose}] - \frac{k_{AR} d[\text{Ribosome}(t)]}{dt} - \frac{k_{AP} d[\text{Protein}(t)]}{dt}$$

where  $k_{AR}$  is the number of amino acid molecules used to form one ribosome molecule ( $k_{AP}$  is similarly defined).

One can solve for  $K'$  using the same procedure as before thus obtaining

$$\text{Amino}(0) = K' \cdot \text{Glucose}(0) / \ln 2 - k_{AR} \cdot \text{Ribosome}(0) - k_{AP} \cdot \text{Protein}(0).$$

In this way rate constants can be calculated for different chemical pools in different environments.

Assume that for system ( $S, \tau$ ) values of  $K(\bar{B})$  and  $\bar{B}$  have been computed for two different environments and that in similar fashion the constants  $k, k_B, k_{BB}$  have also been estimated from the literature. This results in two simultaneous linear equations which can be solved for  $P_1$  and  $P_3$ . Let the values so obtained by  $\bar{P}_1$  and  $\bar{P}_3$  and consider the form of  $K(B)$  derived from equation (6):

$$K(B) = \frac{k + (\bar{P}_1 \cdot k_B)B + (\bar{P}_3 \cdot k_{BB})B^2}{1 + \bar{P}_1 \cdot B + \bar{P}_3 \cdot B^2} \tag{7}$$

Using this function for  $K(B)$  we see that equation (3) will hold and  $h$  will be a homomorphism to the extent that:

- (1) The steady-state conditions assumed in equation (4) are satisfied, and
- (2) The values obtained for  $P_1$  and  $P_3$  in two environments hold for other environments of interest.

The steady-state conditions will hold to the extent that, as is usually assumed, the adjustment in enzyme form concentrations is rapid with respect to the formation of product from substrate. Also, if the values obtained for  $P_1$  and  $P_3$  do not hold for other environments one can complicate system  $S$  by assuming more forms for the enzyme-product complex, thus introducing higher order terms in equations (6) and (7).

We see that in order to fully discuss the notion of homomorphism in practical applications we need to determine how close an arbitrary map comes to being a homomorphism. Ulam (1966) has suggested the introduction of a numerical error function which in this context takes the form  $\varepsilon[\tau'h(s)]$ ,

$h[\tau(s)]$ . Then the extent to which  $h$  is a homomorphism is measured by the least upper bound on  $\varepsilon$  over the region of the state space of interest.

Let us put this example in a more general perspective. We have a system  $S$  in which allosteric inhibition is modelled by the presence and effects of a number of different enzyme-product complexes. This system is not used in our living cell model but is replaced by a simpler system  $S'$ . In  $S'$ , allosteric inhibition is effected by a feedback path from the product to the rate "constant" governing its rate of formation. The reader may wish to check that as far as the substrate and product are concerned the behavior of the two systems is the same (given that equation (3) is valid). *In fact, this kind of behavior preservation is just that entailed by requiring that  $S'$  be a homomorphic image of  $S$ .* We see here how the homomorphism concept elucidates a technique actually used to implement and simplify the allosteric inhibition component of our model.

#### 4. Complexity Reduction of Aggregation

##### (A) COMPUTER SIMULATION COMPLEXITY MEASURES

Investigation of computational complexity is actively being pursued by computer scientists (e.g. ACM Symposium on Theory of Computing, 1969) but little attention has been devoted to measures of complexity relevant to simulation of a model system on a computer. We shall introduce two very basic measures here which are directly related to the representing digraph of a system. One is related to the memory capacity required to store the current state of the system—we require at least  $\log_2 |S|$  bits of storage if there are  $|S|$  states. Thus our first measure is  $memory(S) = \log_2 |S|$ . Since  $S = \prod_{\alpha \in D} S_\alpha$  we

have  $memory(S) = |D| \cdot \left( \frac{1}{|D|} \sum_{\alpha \in D} \log |S_\alpha| \right)$  which indicates the proportionality of the memory capacity to the number of co-ordinates  $|D|$  if all states sets are the same size.

We can also measure the complexity associated with coding the transition function as a computer program. Assume (not unrealistically) that the number of instructions necessary to encode  $\tau_\alpha$  is proportional to the number of co-ordinates in its domain, i.e. to  $|I_\alpha|$ . Then the length of the program for  $\tau$  is proportional to the sum  $\sum_{\alpha \in D} |I_\alpha|$ . Thus our second measure is  $length(D) = \sum_{\alpha \in D} |I_\alpha|$ , which is, in fact, just the number of lines in the digraph. Also, the time to run the program (for one state transition) is "somewhat" proportional to its length. The existence of iterations complicates this relationship.



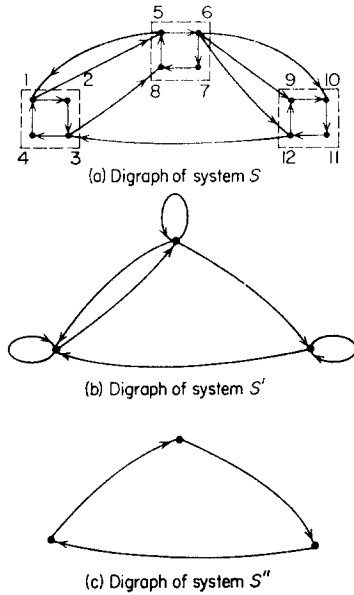
Both *memory* ( $S$ ) and *length* ( $D$ ) cannot exceed the bounds set by the computational capacity of real computers. It is this fact, among others, which would thwart attempts to use the state spaces at the atomic, molecules and concentration levels of Table 1. In other words, were we to try describing the state of a biological cell by listing the states of each of the elementary particles of which the cell is composed (the atomic model), we would hardly have enough data storage capacity to keep track of such a long list. Moreover, we would not have enough program storage capacity to specify how each atomic state changes as a function of the prior states of the atoms which influence it. (In addition, the time required to run such a program would exceed a scientist's patience if not his lifetime). In our new terminology, then, we seek model systems whose memory and length complexities are small enough to enable an actual simulation on a real computer.

#### (B) COMPLEXITY REDUCTION BY AGGREGATION

We shall now give a fairly general mathematical formulation for the process of aggregation or lumping by which relatively simple systems are derived from more complex ones. This kind of process was illustrated in the preceding example and is basic to the model of the *E. coli* cell from which it was taken. [Recall the manner in which the chemicals in the cell were combined into pools (Figs 1 and 2).] We shall relate the aggregation process to the more basic notion of homomorphism, deriving thereby necessary and sufficient local conditions which guarantee that such a process will yield a valid homomorphic image system. At the same time we shall see how the complexity measures introduced before may be simultaneously reduced by aggregation.

Suppose then we are given a system  $(S, \tau)$  with co-ordinates  $D$ , state space  $S = \times_{\alpha \in D} S_\alpha$  and transition function given by  $\{\tau_\alpha | \alpha \in D\}$  where  $\tau_\alpha : \times_{\beta \in I_\alpha} S_\beta \rightarrow S_\alpha$ . We wish to construct a homomorphic image of  $S$  by aggregating co-ordinates. This process can be decomposed into a composition of two sub-processes. In process 1 the co-ordinates are lumped together but the state space is not reduced. In process 2 the co-ordinates of the system are retained but the state space is reduced. Figure 4 illustrates the application of these processes.

For process 1, partition the co-ordinate set  $D$  into blocks  $B_1, B_2, \dots$  (where the blocks are mutually disjoint subsets of  $D$  which together exhaust it). The new system  $(S', \tau')$  will have as co-ordinates  $D' = \{B_i\}$ , the state set of co-ordinate  $B_i$  will be  $S_{B_i} = \times_{\alpha \in B_i} S_\alpha$  and the state space is then  $S' = \times_{B_i \in D'} S_{B_i}$ . Note that  $S' = S$  so that the state space has not really been altered in this



System	No. of co-ordinates	Size of each co-ordinate state set	Size of state space	Storage capacity ( $ D  \log_2  S_2 $ )	Program length (No. lines in digraph)
$S$	$ D  = 12$	$ S_\alpha  = 8$	$ S  = 8^{12}$	$memory(S) = 36$	$length(D) = 19$
$S'$	$ D'  = 3$	$ S'_\alpha  = 8^4$	$ S'  = 8^{12}$	$memory(S') = 36$	$length(D') = 7$
$S''$	$ D''  = 3$	$ S''_\alpha  = 8$	$ S''  = 8^3$	$memory(S'') = 9$	$length(D'') = 3$

FIG. 4. Illustrating the aggregation process. A system  $S$  with 12 co-ordinates each having a state of size 8 is simplified to a system  $S''$  having 3 co-ordinates each having the original state set size (8). The intermediate system  $S'$  is obtained by partitioning the 12 co-ordinates of  $S$  into 3 blocks of 4 co-ordinates each ( $1' = \{1, 2, 3, 4\}$ ,  $2' = \{5, 6, 7, 8\}$ ,  $3' = \{9, 10, 11, 12\}$ ) with state set size  $8^4$ . Finally  $S''$  is obtained from  $S'$  by reducing the state set size of each co-ordinate to 8.  $S''$  is less complex than  $S$  with respect to the memory and length complexity measure.

process. Define the transition function for any co-ordinate  $B_i$  as

$$\tau'_{B_i}(s) = (\tau_{\alpha_1}(s), \tau_{\alpha_2}(s), \dots, \tau_{\alpha_n}(s))$$

for each  $s \in S'$  ( $= S$ ) where  $B_i = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ . One can verify that the subset of  $D'$  on which  $\tau'_{B_i}$  actually depends is  $I'_{B_i} = \{B_j/B_j \cap J_{B_i} \neq \phi\}$  where  $J_{B_i} = \bigcap_{\alpha \in B_i} I_\alpha$ . This is because  $J_{B_i}$  is the set of co-ordinates on which the functions  $\{\tau_\alpha | \alpha \in B_i\}$  depend and  $I'_{B_i}$  is the set of blocks which contain co-ordinates in  $J_{B_i}$ .

It is readily verified that the system  $(S', \tau')$  so defined is isomorphic to the system  $(S, \tau)$ . [Under the identity map  $i : S \rightarrow S', i(s) = s$ ].

Constructing the digraph  $D'$  from the sets  $I'_{B_i}$  we find that there is a line from  $B_i$  to  $B_j$  in  $D'$  just in case there are points  $\alpha$  in  $B_i$  and  $\beta$  in  $B_j$  and there is a line from  $\alpha$  to  $\beta$  in the digraph  $D$ . [This is called the *condensation* of  $D$  with respect to the partition  $D'$  (Harary, Norman & Cartwright, 1965)]. This means that while the *number of points in the digraph may be made to decrease in going from  $D$  to  $D'$ , the total number of lines in the digraph cannot increase* (though it may not decrease).

In process 2 starting with a system  $(S, \tau)$  with co-ordinates  $D$  and state space  $S = \times_{\alpha \in D} S_\alpha$  we construct a system  $(S', \tau')$  also having co-ordinates  $D$  but with (possibly smaller) state space  $S' = \times_{\alpha \in D} S'_\alpha$ . We assume that for each  $\alpha$  there exists an onto map  $h_\alpha : S_\alpha \rightarrow S'_\alpha$  and define the map  $h : S \rightarrow S'$  given by  $h(s_\alpha, s_\beta, s_\gamma, \dots) = [h_\alpha(s_\alpha), h_\beta(s_\beta), h_\gamma(s_\gamma), \dots]$  for each  $(s_\alpha, s_\beta, s_\gamma, \dots) \in S$ . We seek a condition which makes  $h$  a homomorphism. The condition, which is both necessary and sufficient, is that for each  $\alpha \in D$ ,

$$h_\alpha[\tau_\alpha(s_{\beta_1}, \dots, s_{\beta_n})] = \tau'_\alpha[h_{\beta_1}(s_{\beta_1}), \dots, h_{\beta_n}(s_{\beta_n})]$$

for all  $(s_{\beta_1}, \dots, s_{\beta_n}) \in \times_{\beta \in I_\alpha} S_\beta$ . We call this the condition of *preservation of co-ordinate functionality* since it translates the *global* requirements for  $S'$  to be a homomorphic image of  $S$  into *local* requirements involving the transition functions  $\tau_\alpha$  and  $\tau'_\alpha$  at each co-ordinate  $\alpha$ . In particular if for any  $\alpha$  a function  $\tau'_\alpha$  can be made to satisfy the condition then it *depends at most on the co-ordinates  $I_\alpha$* . In fact if any simplification is achieved  $I'_\alpha$  will be a proper subset of  $I_\alpha$ . This means that  $D'$  is a *subdigraph* of  $D$ , i.e. if a line from  $\alpha$  to  $\beta$  appears in  $D'$  it must also appear in  $D$ . Thus in this case, in going from  $S$  to  $S'$  the *total number of lines may be made to decrease* (it cannot increase).

To establish the condition, we use the maps,  $\text{proj}_\alpha : S \rightarrow S_\alpha$  where  $\text{proj}_\alpha(s)$  is the projection of  $s$  on the  $\alpha$  co-ordinate. Thus  $\text{proj}_\alpha[h(s)] = h_\alpha(s)$  and  $\text{proj}_\alpha[\tau(s)] = \tau_\alpha(s)$  for each  $\alpha \in D$ . Then  $\tau'[h(s)] = h[\tau(s)]$  if, and only if,  $\text{proj}_\alpha[\tau'(h(s))] = \text{proj}_\alpha[h(\tau(s))]$  for each  $\alpha \in D$ . Now  $\text{proj}_\alpha[\tau'(h(s))] = \tau'_\alpha(h(s))$  and  $\text{proj}_\alpha[h(\tau(s))] = h_\alpha(\tau_\alpha(s))$  so that  $h$  is a homomorphism if, and only if  $\tau'_\alpha(h(s)) = h_\alpha(\tau_\alpha(s))$ . Now  $h_\alpha(\tau_\alpha(\cdot))$  is a function of at most the co-ordinates in  $I_\alpha$  and for the equality to hold for *all*  $s \in S$  it can be readily seen that  $\tau'_\alpha(h(\cdot))$  is necessarily a function of these same co-ordinates. This establishes the necessity of the condition as given before. Its sufficiency is easily verified. [If we demand that the homomorphism hold for only a subset of  $S$ , a similar result follows but it must be more carefully phrased.]

Suppose that starting with a system  $(S, \tau)$  we apply process 1 to arrive at system  $(S', \tau')$  with fewer co-ordinates and then apply process 2 to  $(S', \tau')$

to arrive at a system  $(S'', \tau'')$  with a smaller state space and whose digraph  $D''$  has fewer lines than  $D$ . Because of its smaller state space  $memory(S'') < memory(S)$  and because it has fewer lines  $length(D'') < length(D)$ . The net result is that  $(S'', \tau'')$  is a homomorphic image of  $(S, \tau)$  whose complexity, as determined by the memory and length complexity measures, is decreased.

Note that the digraph  $D''$  is a subdigraph of a condensation of the digraph  $D$ . This means that not only has the state behavior been preserved in going from  $(S, \tau)$  to  $(S'', \tau'')$  but also the *local structure*, i.e. the manner in which the co-ordinates are interdependent. This preservation of local structure is *not* a necessary consequence of the preservation of state behavior. In general, a system  $(S', \tau')$  may be a homomorphic image of a system  $(S, \tau)$  while  $D'$  is not a subdigraph of a condensation of  $D$ . The preservation of local structure *is*, however, a consequence of the aggregation process and the requirements for homomorphism.

## 5. Discussion

In this paper, we have utilized system theory concepts to describe and analyze the basic character of our simulation of a living cell. In contrast to other simulation of cell metabolism, we simplified our model by lumping the various molecular species into pools and then attempted to simulate the behavior of these pools over time. This aggregation technique greatly reduced the complexity of our simulation but raised the issue of the validity of our model. In this regard, we gave necessary and sufficient conditions which must obtain if an aggregated system is to model (in the sense of being an exact homomorphic image) the original system. These conditions were used to show that the memory and length measures of complexity can be simultaneously reduced in a valid aggregated model.

In contrast to models in general, aggregated models preserve the local structure of the co-ordinate dependency digraph. This is important since the greater is this preservation of structure the more readily can components of the model be identified with corresponding components of the real world biological system and their behavior compared. [The identification of components, i.e. systems formed by subsets the co-ordinate set, can be accomplished easily when the co-ordinates of the system and a model are related by a (univalent) mapping as for aggregated models but is much more difficult in the more general case where the relation between co-ordinates may be many to many.] As we have indicated, the possibility of checking the behavior of components of our model of the cell against that of corresponding real world components played a significant role in determining the assignment of molecules to pools.

## 6. Conclusions

System theory concepts help to better understand and formulate problems encountered by biologists and other scientists in modeling and simulation. At the conceptual level, these concepts help explicate the notions of model complexity and adequacy. At the more practical level they may suggest and justify techniques for the reduction of complexity, and for the testing of model adequacy.

In particular the aggregation of entities as a modeling technique has been given a system theoretic formulation. This technique is fundamental to our model of the living cell and may provide a powerful tool for other simulation studies.

This work was supported by an Institute of Science and Technology Fellowship awarded to one of us (B. P. Z.) and by a Public Health Service Fellowship 2-F3-CA-6931-03 from the National Cancer Institute awarded to the other of us (R. W.); Public Health Service Grant GM-12236 from the National Institute of Health, Army Research Office—Durham Contract DA-31-124-ARO-D-483 and Office of Naval Research Contract N00014-67-A-0181-0011 and the National Science Foundation Grant No. GJ-519.

## REFERENCES

- ACM Symposium on Theory of Computing (1969). Marina del Rey, Calif.
- ARBIB, M. A. (1969). *Theories of Abstract Automata*. Englewood Cliffs, N.J.: Prentice Hall.
- CHANCE, B., ESTABROOK, R. W. & WILLIAMSON, J. R. (1965). *Control of Energy Metabolism*. New York: Academic Press.
- FEIN, L. (1966). *Natural Automata and Useful Simulations*, p. 181. New York: Spartan.
- GARFINKEL, D. (1966). *J. biol. Chem.* **241**, 286.
- GREEN, H. A. J. (1964). *Aggregation in Economic Analysis*. Princeton: Princeton University Press.
- GOODWIN, B. (1969). In *Towards a Theoretical Biology*, Vol. 2. (Waddington, C. H., ed.) Chicago: Aldine.
- HARARY, F., NORMAN, D. & CARTWRIGHT, R. (1965). *Structural Models: An Introduction to the Theory of Directed Graphs*. New York: John Wiley.
- HARTMANIS, J. & STEARNS, R. E. (1966). *Algebraic Structure Theory of Sequential Machines*. Englewood Cliffs, N.J.: Prentice Hall.
- HEINMETS, F. (1966). *Analysis of Normal and Abnormal Cell Growth*. New York: Plenum Press.
- IRIJI, Y. (1968). *Econometrica* **36**, 252.
- KALMAN, R. E., FALB, P. L. & ARBIB, M. A. (1969). *Topics in Mathematical Systems*. New York: McGraw-Hill.
- KAUFFMAN, S. A. (1969). *J. theor. Biol.* **22**, 437.
- KLIR, J. & VALACH, M. (1967). *Cybernetic Modelling*. Princeton: Van Nostrand.
- KOCH A. I. (1967). *J. theor. Biol.* **15**, 75.
- KROHN, K. & RHODES, J. (1965). *Trans. Am. math. Soc.* **116**, 4.
- ROOSEN RUNGE, P. (1967). *Math. Systems Theory*, **2**, 203.
- ROSEN, R. (1968). *J. theor. Biol.* **18**, 380.

- SAVAGEAU, M. A. (1969). *J. theor. Biol.* **25**, 365.
- SIMON, H. A. & ANDO, H. (1961). *Econometrica*, **29**, 111.
- SIMON, H. A. (1962). *Proc. Am. phil. Soc.* **106**, 6.
- ULAM, S. M. (1966). In *Prospects for Simulation and Simulators of Dynamic Systems* (Shapiro, G. & Rogers, M., eds). New York: Spartan.
- WEINBERG, R. (1968a). *Bact. Proc.* **114**, Abs G 114.
- WEINBERG, R. (1968b). *Genetics*, **60**, 235.
- WEINBERG, R. & BERKUS, M. (1969a). *Biometrics*, **25**, 447.
- WEINBERG, R. & BERKUS, M. (1969b). *Univ. Michigan tech. Rep. 01252-2-T*. Ann Arbor, Michigan.
- WEINBERG, R. & ZEIGLER, B. P. (1970). *Communs Am. Soc. Cybernetics* (in the press).
- YEISLEY, W. G. & POLLARD, E. C. (1964). *J. theor. Biol.* **7**, 485.
- ZEIGLER, B. P. (1968). Doctoral dissertation. University of Michigan.
- ZEIGLER, B. P. (1969). *Proc. 3rd a. Princeton Conf. Systems Inf. Sci.* Princeton: Princeton University Press.