J. theor. Biol. (1991) **152,** 429-453

A Connectionist Model of Development

ERIC MJOLSNESS†‖, DAVID H. SHARP‡¶ AND JOHN REINITZ§††

† *Department of Computer Science, Yale University, P.O. Box* 2158 *Yale Station, New Haven, CT* 06520-2158, ‡ *Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM* 87545, § *Department of Biological Sciences, Columbia University, New York, NY* 10027 *and* ‡*Center for Medical Informatics, Department of Anesthesiology, Yale Medical School,* 333 *Cedar St, New Haven, CT* 06510‡‡, *U.S.A.*

(*Received on* 27 *June* 1990, *Accepted in revised form on* 30 *January* 1991)

We present a phenomenological modeling framework for development. Our purpose is to provide a systematic method for discovering and expressing correlations in experimental data on gene expression and other developmental processes. The modeling framework is based on a connectionist or "neural net" dynamics for biochemical regulators, coupled to "grammatical rules" which describe certain features of the birth, growth, and death of cells, synapses and other biological entities. We outline how spatial geometry can be included, although this part of the model is not complete. As an example of the application of our results to a specific biological system, we show in detail how to derive a rigorously testable model of the network of segmentation genes operating in the blastoderm of *Drosophila*. To further illustrate our methods, we sketch how they could be applied to two other important developmental processes: cell cycle control and cell-cell induction. We also present a simple biochemical model leading to our assumed connectionist dynamics which shows that the dynamics used is at least compatible with known chemical mechanisms.

1. Introduction

In the course of growth, living cells undergo a variety of transitions: they can divide to form more cells of the same or different type; they can produce fibers which form synapses and contacts; and they can interact in various ways with nearby cells. During these processes, the internal state of the cells may be changing. We believe that it is possible to answer important questions about development without modeling all aspects of the growth process at an equal level of detail. For example, we think it plausible that important features of genetic regulatory circuits can be understood in a model which treats mitosis as an elementary event. Moreover, we believe that a phenomenological model of gene regulation can be inferred from observational data without a full understanding of the underlying biochemistry. In line with these ideas, we introduce in this paper a modeling framework for describing some of the principal processes which occur in development. We describe a specialization of

‖ Supported in part by the Air Force Office of Scientific Research under grant AFOSR 88-0240.
¶ Work supported by the United States Department of Energy.
†† Supported in part by grant 5-T15-LM07056 from the National Library of Medicine.
‡‡ Present address and address to which correspondence should be sent.

0022-5193/91/200429+25 $03.00/0
© 1991 Academic Press Limited

this framework to model the network of gene regulation occurring in the blastoderm of *Drosophila*, for which considerable gene expression data is available. Initial results from simulations of our model are compared to data in Reinitz *et al.* (1991). In this paper we also show how the framework could be applied to model cell cycle control and cell-cell induction, and we present a simple biochemical model of the dynamics we assume for the genetic regulatory circuit, to show their compatibility with known chemical mechanisms.

A configuration of a developing embryo is specified by the number and internal state of fundamental objects such as cells, nuclei, fibers and synapses. We wish to model the transition between two such configurations. The model must describe two kinds of processes: the change in internal state of the objects, and the discrete change in the number and kind of objects. Our model involves selecting which kind of process operates at each time for each object, the possible processes being specified by the production rules of a recursive grammar, and what the internal state of the resulting objects are, as determined by a connectionist model analogous to a neural net. Since spatial information is as important as genetics in development, we also model the effect of cell-cell adjacency by a connectionist network, and we discuss mechanical motion briefly.

2. A Phenomenological Modeling Framework

Suppose we have a single population of primitive objects such as cells, nuclei, fibers, and synapses. We represent the internal state of each object i in this population by a vector $\boldsymbol{v}_i$. The components of $\boldsymbol{v}_i$ are denoted v_i^a, indexed by a. For example, the internal state of a cell nucleus might be given by specifying the concentrations of a certain set of transcription factors; that of a synapse by membrane voltage, internal Ca^{2+} concentration, and a set of ionic conductances. To illustrate our approach, we focus in this section on the case of a cell nucleus. A more general formulation which can describe a variety of interacting objects is given in section 5.

A nucleus undergoes two classes of state transitions. During interphase its genes are synthetically active and regulate one another; during mitosis the nucleus duplicates while keeping its genome silent. The result is a pattern of regulatory connections between the gene products. We represent the interaction of a pair of genes a and b by a single real number T^{ab}. This corresponds to a "synaptic strength" (or "connection strength") in a "neutral network" (or "connectionist") model. The collection of such numbers forms a connection matrix **T**. Characterizing the interaction between two genes by one real number captures a notion commonly used by geneticists that one gene activates or represses another (positive or negative connection strengths, respectively), or that two genes do not interact (connection strength of zero). In addition, it allows for differing strengths of activation or repression between genes. This idea is an extension of the observation (Muller, 1932) that most of the alleles at a genetic locus can be ordered according to a scalar measure of function, and that alleles that do not fall on this one-dimensional scale (neomorphs) are rare.

There may be many proteins b regulating the gene for protein a and thus influencing the dynamics of v_i^a. To make a tractable model, we will assume that these effects are monotonic in the concentrations v_i^b and are approximately additive, with non-linearity confined to sigmoidal threshold functions g_a. So we assume that the dynamics of v_i^a depend on the other variables v_i^b through a summed input u_i^a:

$$u_i^a = \sum_b T^{ab} v_i^b + h^a, \tag{1}$$

where h^a determines the threshold of g_a. During the time when object i does not participate in birth or death processes, which for dividing cells is called interphase, we use a continuous time model of its internal dynamics. A simple model is the neural net dynamics given by [cf. (Hopfield, 1984)]

$$\tau_a \frac{dv_i^a}{dt} = g_a(u_i^a) - \lambda_a v_i^a. \tag{2}$$

The long-time behavior of this system has been studied, and in some cases is controlled by simple limit sets such as fixed points [$v = g(u)/\lambda$]. In specific cases there are further arguments justifying this form of the equation. In section 4 we discuss a slight modification of eqn (2) which follows from a highly simplified biochemical model of regulatory interaction via binding sites associated with each of the genes in a genetic circuit.

Equations (1) and (2) comprise our connectionist model of the internal dynamics of interphase, one of many processes involved in development. Other continuous time and purely internal dynamical subsystems will be modeled in like manner, with just the connection matrix $\mathbf{T}$ and the thresholds h^a changed.

The connection matrix elements T^{ab} and the threshold terms h^a are regarded as adjustable parameters, to be fit to biological data on patterns of gene expression or other observable biochemical state variables. Such parameter fitting is generally called "training" in the field of neural network learning, and has been extensively studied. Consequently there are many algorithms to perform the fit (Rummelhart *et al.*, 1986; Mjolsness *et al.*, 1989; Pearmutter, 1989) and several theoretical criteria to determine when a good fit to existing data is expected to "generalize" and agree with further data, as yet unseen by the training algorithm (Baum & Haussler, 1989). In particular these issues were studied in the context of the recursive growth of synthetic neural networks, under the constraint of parsimonious network description, in Mjolsness *et al.* (1989).

In many situations the state variables (v_i^a) are observable and it is possible to perturb the system by an experimental manipulation that inactivates the chemical species represented by a given state variable. In the case of gene products, a state variable is inactivated by mutating its gene; manipulations for other types of state variables include the administration of specifically acting drugs and toxins. In these cases it will be possible to unambiguously compare the model to data. For a network of N interacting state variables, the circuitry will be represented by a matrix of real numbers having N^2 elements. For a given cell only the N state variables may be observed, which is not enough information from which to obtain the N^2 matrix

elements. If we consider an ensemble of experimentally perturbed systems, each of which contains one or more inactivated state variables, the problem disappears. The key observation is that a cell with a single inactivated state variable is represented by a connection matrix that has one row set to zero. This alters the dynamics, and hence the observed values of the N state variables. Consideration of the unperturbed system in conjunction with the N perturbed systems, each of which has a single inactivated state variable, shows that there will be N^2 pieces of data for N^2 unknowns. Because it is often practical to inactivate two or more state variables at once, additional observational constraints on the model are available.

We next consider birth processes such as cell division (mitosis). For some purposes, we regard such processes as elementary events, which can be modeled with a discrete time neural net update equation with its own connection matrix. Since the different daughter cells of one parent cell are not necessarily equivalent, we use one such connection matrix $\mathbf{T}_k$ (with components T_k^{ab}) for each of the progeny. We use multiple index notation: if i is the index of the parent cell, then (i, k) is the index of its kth daughter cell. [And $((i, k), l)$ would index the second generation descended from i.] We then suppose that

$$v^a_{(i,k)} = v^a_i + R_a g_a \left(\sum_b T^{ab}_k v^b_i + h^a \right). \tag{3}$$

This is a phenomenological model in which a birth process is modeled by an update of a discrete time neural network. A simplified, linear version of this model was proposed in section 4.1 of Mjolsness *et al.* (1990). An important special case is obtained if $\mathbf{T}_k$ is diagonal, and if g is linear and h is zero for all a. Then we can specialize eqn (3) to describe an important biological situation: unequal partitioning of gene products among the progeny of a parent cell, with no further production, gene regulation, or other genetic computation during mitosis:

$$v^a_{(i,k)} = U^a_k v^a_i, \tag{4}$$

where $U^a_k \geq 0$.

For a system which includes both interphase and cell division, eqns (2) and (3) must be combined so that (2) operates continuously except at certain discrete times when (3) is invoked. The same is true of any combination of continuous and discrete time processes. After such an event, continuous dynamics (2) resumes for the newly generated objects. This is shown in Fig. 1, a schematic illustration of the evolution of a system that undergoes interphase and cell division (in addition to a discrete type change transition, discussed in section 5).

Other processes which may be modeled in this way include axon sprouting and growth (with an internal state vector for each axon segment), synapse formation, and cell-cell induction. These processes each have a rule by which one or two objects may be updated or replaced by several others. Processes that involve the elimination of an object, or the interaction of two objects, will be modeled in section 5.1. Every developmental process that we model has a rule stating which types of object are present at the beginning and the end of the process, and an internal dynamics model like eqns (2) or (3). The set of such rules form a *grammar* Γ in

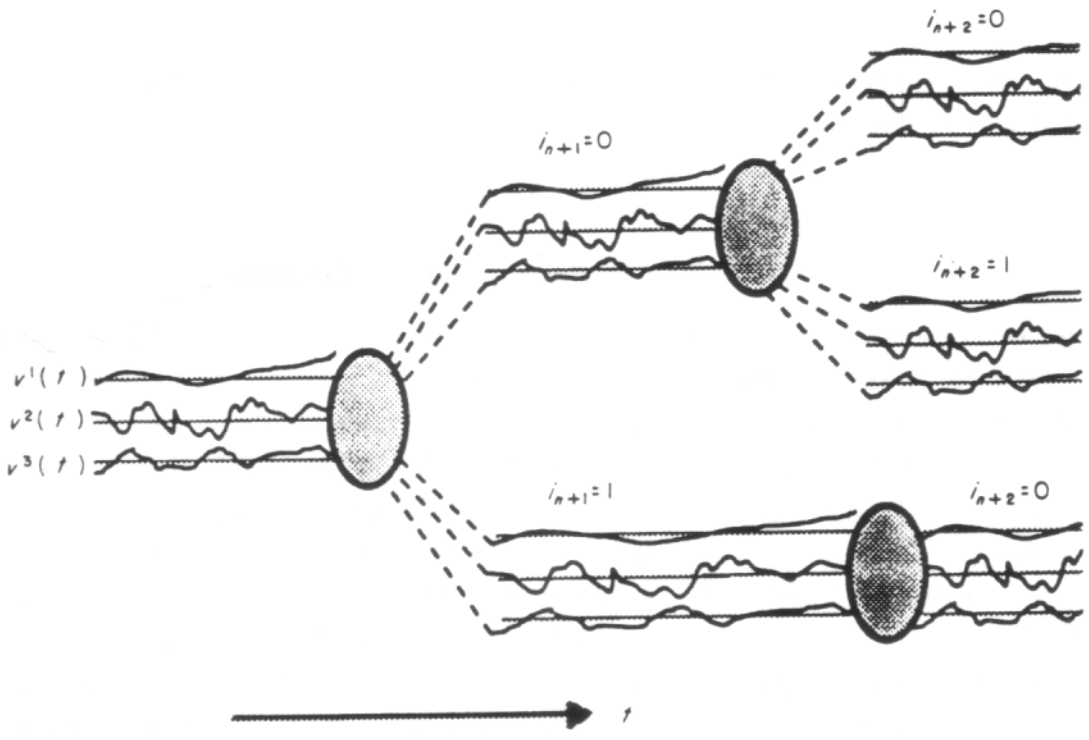

FIG. 1. A schematic illustration of the history of the state variables v_i^a under a combination of continuous and discrete time grammatical rules. Time increases to the right; three generations of objects are shown. Continuous time dynamics is denoted by the stippled horizontal axes; discrete time dynamics by the stippled ovals. The solid black curves are graphs of the functions $v_i^a(t)$. The extreme right oval denotes a type change; the other two are mitoses. The level of a state variable may change under the action of a discrete time rule. Note, for example, that $v^1(t)$ is always above baseline before a mitosis and below it afterwards. The daughter objects' indices are i_{n+1} and i_{n+2} depending on generation number; in section 2 these are denoted k and l respectively.

the manner of the Lindenmayer Systems (Lindenmayer, 1968). By a grammar, we mean a set of rules by which a single object of a given type may be replaced by one or more objects of the same or different types. The simplest grammars leading to synapse formation, for example, would include type transitions among cells, fibers, and synapses, each containing heritable state vectors. In general, the grammatical rule adopted by an object i at a given time t will be a function of its state vector $\boldsymbol{v}_i(t)$, but note that we do not exclude the possibility that in some circumstances the rule choice will be driven by an outside forcing term.

When augmented with geometric dynamics based on motion of cells and fibers through a medium, we arrive at a class of models with three interacting sectors: *internal dynamics*, *grammar*, and *geometry*. The result is a schematic framework for modeling development.

In section 5, we present two classes of grammatical rules. One class contains continuous internal dynamics; the other does not. A given state transition may or may not be represented by a rule with internal dynamics; the choice depends on the specific biological system at hand and the nature and quality of experimental data available. For example, a change of cell type could be modeled by a discrete time grammatical rule, or it could emerge from a network of genes settling to a stable-state of expression. In the next section we use the second type of formulation in a case where there is rich data on gene expression; but in section 6 we use a

discrete time grammatical rule to model a cell type transition during induction in a system where there is presently little gene expression data.

3. Application to the *Drosophila* Blastoderm

As a concrete example of the preceding formalism, we describe its application to an exceptionally simple biological system, the blastoderm of *Drosophila melanogaster*. Immediately following fertilization, the zygotic nuclei undergo a rapid series of mitoses without the formation of cells. After eight almost synchronous divisions, these nuclei migrate to the cortex of the egg, whereupon transcription of the zygotic genes begins. This stage is called the syncytial blastoderm, because no cells are present. After another five divisions, cell membranes are layed down and gastrulation begins (Foe & Alberts, 1983). The timing of these cell divisions is under the control of maternal gene products. The protein products of pattern formation genes essential for laying down the basic body plan of the animal are expressed at this time in patterns that rapidly evolve from coarse to fine scale spatial resolution [reviewed in Akam (1987) and Ingham (1988)].

The *Drosophila* egg is approximately an ellipsoid, but asymmetries in its shape clearly define two axes, each with a polarity. These axes provide co-ordinates for the blastoderm as well. One axis runs in an anterior-posterior direction, and the other in a dorsal-ventral direction. The pattern formation genes fall into two classes. To a reasonable degree of approximation, the level of expression of a member of the first class of genes is solely a function of location on the anterior-posterior axis; these genes are members of the anterior-posterior class. The expression level of a member of the second class of genes depends only on position along the dorsal-ventral axis; these genes belong to the dorsal-ventral class.

The separation of these two classes of genes by expression pattern carries over to their dynamical interactions. A member of one of these classes of genes does not regulate the expression of a member of the other class during the blastoderm stage, except perhaps in the region of the anterior or posterior pole. For the rest of this section we focus on the zygotic anterior-posterior pattern formation genes, often referred to as segmentation genes. The segmentation genes are dynamically coupled in a network of genetic regulation. A line of evidence leading to this conclusion is the observation that disabling one segmentation gene by mutation leading to this conclusion is the observation that disabling one segmentation gene by mutation causes changes in the pattern of expression of many of the other segmentation genes (Carroll & Scott, 1986; Jackle *et al.*, 1986; Rushlow *et al.*, 1987; Frasch & Levine, 1987; Carroll *et al.*, 1988; Pankratz *et al.*, 1989; Reinitz & Levine, 1990). The characterization of this regulatory network is one of the objectives of our modeling effort. A precise formulation of the regulatory network is required to interpret altered patterns of gene expression in terms of regulatory action. We call attention to related studies of this problem (Meinhardt, 1986; Lacalli *et al.*, 1988; Goodwin & Kauffman, 1990); our work differs in its emphasis on the regulatory circuitry. An approach closer in spirit to ours is given in recent work of Edgar *et al.* (1989).

We next describe how eqns (1), (2) and (3) are applied to nuclei in the blastoderm. Since we are only considering segmentation genes, we can approximate the blastoderm by a line of equally spaced nuclei along the anterior–posterior axis, indexed by i in such a way that nucleus $i+1$ is immediately posterior to nucleus i. Each nucleus i contains an internal state vector v_i^a whose elements are N protein concentration indexed by a. In this problem we use four grammatical rules. The first rule describes mitosis. The second rule governs the change in internal state during interphase; that is, the dynamics of gene regulation. Two additional rules govern spatial interactions; there is one such rule for each of the nucleus is two nearest neighbors. We allow the last three rules, which all prescribe continuous internal dynamics, to operate simultaneously. The operation of these three rules and the mitosis rule are mutually exclusive. Which grammatical rule to use at a given time is set maternally, rather than being determined by the state vector $\boldsymbol{v}_i$.

We first consider mitosis. Nuclei in the blastoderm appear to undergo symmetric mitoses, with equal partitioning of protein products in the two daughters. Moreover, no synthesis of gene products takes place during mitosis. The grammatical rule that represents this process makes two copies of a nucleus i, preserving the state vector v_i^a. It is easy to incorporate this rule choice into the dynamical calculation, since it amounts to utilizing eqn (4) with U equal to unity.

Consider now the remaining two rules, one of which concerns spatial interaction. We regard protein synthesis as a unitary process, and do not assign a separate variable to the concentration of mRNA synthesized from the ath gene. Such a model is a lumped description of both transcriptional and translational control. If these nuclei were isolated from one another, we could describe the time evolution of the concentrations of segmentation gene products by eqns (1) and (2), with $1/\tau_a = R_a$, the maximum rate of synthesis from gene a. The situation is in fact more complicated, because there are no cell membranes in the syncytial blastoderm and nuclei may exchange material. This can still be treated using eqn (2), but we must modify eqn (1) to reflect the added spatial interaction.

First, introduce

$$\hat{u}_i^a = \sum_b T^{ab} v_i^b + h^a. \tag{5}$$

The corresponding concentration variable may be called w_i^a and evolves according to

$$\tau_w \frac{\mathrm{d}w_i^a}{\mathrm{d}t} + w_i^a = g_a(\hat{u}_i^a). \tag{6}$$

Let τ_w be very small compared to any other time scales in the network, so that on the time scale with which the v variables move, each w adapts instantly to the fixed point value appropriate to the current v variables:

$$w_i^a = g_a\left(\sum_b T^{ab} v_i^b + h^a\right). \tag{7}$$

To write eqn (1) in a form that allows a diffusion-like exchange of gene products, set

$$u_i^a = \sum_j \Lambda_{ij} v_j^a + w_i^a, \tag{8}$$

where Λ_{ij} is a matrix which describes a chemical coupling between blastoderm nucleus i and other blastoderm nuclei j, and the w_i^a term specifies a direct input from w_i^a to the corresponding v_i^a. We take Λ_{ij} to be of the form

$$\Lambda_{ij} = \frac{D}{l^2}[(\delta_{i-1j} - \delta_{ij}) + (\delta_{i+1j} - \delta_{ij})], \tag{9}$$

where the parameter D is to be fit to data, and δ_{ij} is the Kronecker delta. Note the presence of two terms, one for each nearest neighbor. Spatial geometry enters explicitly through l, the distance between a neighboring pair of nuclei. Every time a synchronous nuclear division takes place, l will decrease by a factor of two. Hence, we may write $D/l^2 = D(n)$, where n is the number of nuclear divisions that have taken place, and $D(n) = 4D(n-1)$. $D(0) = D$, the adjustable parameter of eqn (9).

When substituting eqn (5) into eqn (2) we take $g_a = g_{\text{lin}}$ for all a, where g_{lin} is a special thresholding function equal to a line with unit slope throughout the dynamic range of u. We expand eqn (2) in terms of this definition of u to obtain explicit dynamical equations for interphase:

$$\frac{dv_i^a}{dt} = R_a g_a\left(\sum_{b=1}^{N} T^{ab} v_i^b + h^a\right) + D(n)[(v_{i-1}^a - v_i^a) + (v_{i+1}^a - v_i^a)] - \lambda_a v_i^a. \tag{10}$$

Note that the simultaneous application of three rules is expressed by placing a term for each rule into the expression for u_i^a. This is an illustration of our method for treating multiple rules, the general formulation of which is presented in section 5.

4. The Connectionist Model from a Molecular Viewpoint

In this section we show how molecular and thermodynamic arguments applied to biochemical reactions could lead to a "neural net" equation similar to eqn (2). Figure 2 illustrates the reactions considered. A solution contains several species of

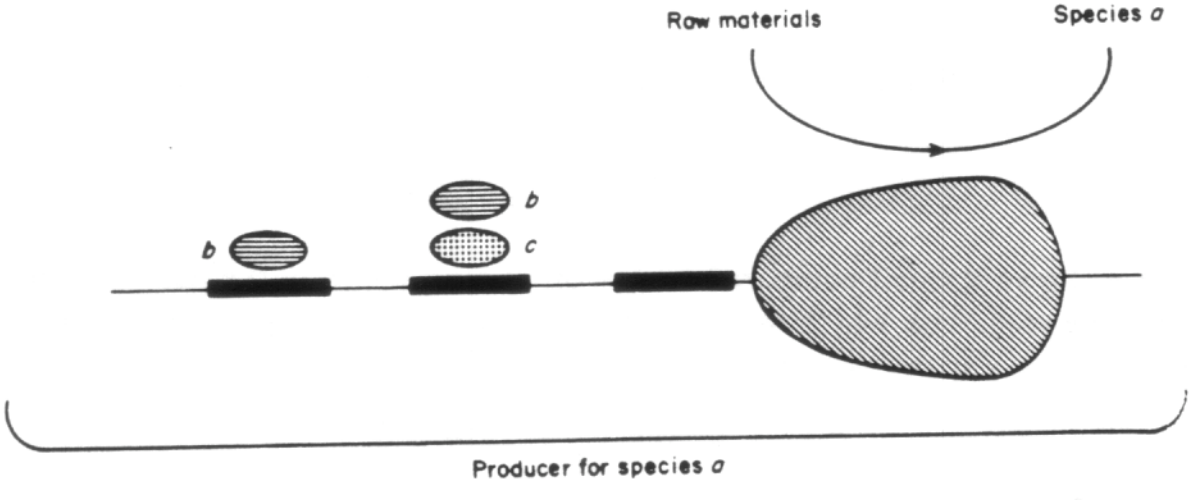

FIG. 2. An illustration of one possible microscopic realization of the biochemical model presented in this section. The producer for species a is represented as a stable transcription complex (diagonal stripes) bound to DNA (horizontal line). The DNA contains enhancer/silencer regions (heavy bars); these may be occupied by other species b and c.

protein. The concentration of species a is $v^a = [a]$. At much lower and fixed concentration $[P^a]$ are the molecules responsible for continuing production of species a. In fact we assume that the concentration of *any* producer molecule is much smaller than the concentration of species a. Each producer molecule has an associated set of binding sites which can influence whether it is in a productive or non-productive state. The occupancy of binding sites is a mechanism by which one gene can activate or repress another.

The processes illustrated in Fig. 2 occur on three distinct time scales. The slowest time scale, here treated as infinitely long, governs the concentration of producer molecules for each species as well as the concentrations of raw materials available for synthesis of new proteins and the concentrations of the decay products of these proteins. For present purposes these are all held fixed. At an intermediate time scale, the concentration of species a is increased by the action of its producer molecules, and depleted as a result of chemical reactions not involving other modeled species:

$$\frac{dv^a}{dt} = R_a \frac{[P^a \text{ in producing state}]}{[P^a]} - \lambda'_a v^a. \tag{11}$$

Two processes occur at a fast time scale. The producer molecule P^a switches between producing and non-producing states in a manner which is correlated with the occupancy of its binding sites. Also, each of m_a binding sites can independently change its state of occupancy by exchanging proteins with the solution. Each binding site can be unoccupied, occupied by a single molecule of species b, or occupied by one molecule of species b together with one molecule of species c. These two processes are assumed to occur fast enough that they are in thermal equilibrium at the intermediate time scale. This thermal equilibrium is described by the grand partition function [generalized from Hill (1985)]

$$\xi_a = K_a x_a^{m_a} + y_a^{m_a}, \tag{12}$$

where

$$x_a = 1 + \sum_b T^{ab} v^b + \sum_{bc} T^{abc} v^b v^c, \tag{13}$$

and

$$y_a = 1 + \sum_b \hat{T}^{ab} v^b + \sum_{bc} \hat{T}^{abc} v^b v^c. \tag{14}$$

In the above equations K_a is the equilibrium constant for the transition between producing and non-producing states of P^a. T^{ab} is the binding constant for species b to one of the binding sites on the producer molecule for species a when it is productive. $\hat{T}^{ab}$ is the corresponding binding constant when this producer molecule is non-productive. To model double occupancy of a binding site by species b and c, one introduces non-zero binding constants T^{abc}. Zero binding constants correspond to prohibited configurations, and there may be many such zero elements in a sparse matrix T. All the binding constants T, equilibrium constants K, and concentrations v are positive or zero. If $T^{ab} > \hat{T}^{ab}$ then b "activates" the production of a, and if $T^{ab} < \hat{T}^{ab}$ then b "represses" the production of a.

We are now able to compute the fraction δ_a of as producer molecules which are in the producing state:

$$\delta_a(\boldsymbol{v}) = \frac{[P^a \text{ in producing state}]}{[P^a]} = \frac{K_a x_a^{m_a}}{K_a x_a^{m_a} + y_a^{m_a}} = \frac{K_a (x_a / y_a)^{m_a}}{1 + K_a (x_a / y_a)^{m_a}}. \tag{15}$$

This equation applies at the same intermediate time scale as eqn (11) and defines a dynamical system. As we shall see, it can be interpreted as a "neural network" model. To place it in a biochemical perspective, we point out that the related equation

$$\delta(c) = \frac{Kc^m}{1 + Kc^m}, \tag{16}$$

is well-known as the "Hill equation" in the special case that c is a single concentration variable. The Hill equation is often used to model co-operative binding systems, an application in which it has a limited domain of validity (Hill, 1985). Our eqn (15) is not such an application since the binding sites on a producer molecule are independent and c is a ratio of low-order polynomials rather than a single variable. Both the Hill equation and eqn (15) must be modified if one generalizes from an equilibrium to a steady-state model, which we will not do here.

In the absence of all repression (i.e. if $\hat{T}^{ab} = 0$ and $\hat{T}^{abc} = 0$) the connection to neural net models is especially clear. No repression implies $y_a = 1$, and therefore

$$\delta_a = g_a(x_a) = \frac{K_a x_a^{m_a}}{1 + K_a x_a^{m_a}}. \tag{17}$$

Equation (11) is simplified to

$$\tau_a \frac{\mathrm{d}v^a}{\mathrm{d}t} = g_a\left(1 + \sum_b T^{ab} v^b + \sum_{bc} T^{abc} v^b v^c\right) - \lambda_a v^a, \tag{18}$$

(where $\tau_a = 1/R_a$ and $\lambda_a = \lambda'_a / R^a$) which is a minor variation on a standard equation of motion for an analog neural network, generalized from Hopfield (1984) by adding T^{abc} terms. With $u^a = g^{-1}(\lambda_a v_a)$, this equation has the form

$$\tau_a \frac{\mathrm{d}u^a}{\mathrm{d}t} + u^a = 1 + \sum_b T^{ab} v^b + \sum_{bc} T^{abc} v^b v^c. \tag{19}$$

The two sets of neural net equations have the same fixed points and differ only by whether the input, u, or the output, v, decays exponentially to its fixed point value. Unlike most neural network models, the network of eqn (18) has no negative connections; such connections require that the form of the differential equation be modified as follows [from eqns (11), (13), (14) and (15)]:

$$\tau_a \frac{\mathrm{d}v^a}{\mathrm{d}t} = g_a\left(\frac{1 + \sum_b T^{ab} v^b + \sum_{bc} T^{abc} v^b v^c}{1 + \sum_b \hat{T}^{ab} v^b + \sum_{bc} \hat{T}^{abc} v^b v^c}\right) - \lambda_a v^a. \tag{20}$$

Equation (20) is our connectionist model. Note that a positive $\hat{T}^{ab}$ in this equation has a similar effect on the argument of g_a as would a negative connection strength T^{ab} in eqn (18). For small values of $\sum_b \hat{T}^{ab} v^b$ and $\sum_{bc} \hat{T}^{abc} v^b v^c$ the two alternatives are indistinguishable.

We have assumed that the different producer molecules act independently during equilibration at the fast time scale. Their only interaction at that time scale is a competition for binding site occupancy from a common solution. This interaction is negligible because binding to each species of producer molecule has negligible effect on the solution. That is because there are many fewer binding sites than potential binding site occupants ($[P^a] \ll [b]$) so that even when all binding sites are filled, there is no change in gene product concentrations $[b]$.

In order to make the picture we have presented a little more concrete, as illustrated in Fig. 2, we might identify the producer molecules P^a with transcription complexes attached to the promoter of the gene coding for gene product a. The binding sites would correspond to enhancers and/or silencers associated with the gene for a. Control of protein synthesis at the level of the translation of mRNA to protein is not included in this picture. The protein species a are participants in a network or circuit of gene regulation, consisting of activation, repression, production, and decay processes.

5. Recursion Equations

The formalism presented in section 2 can be developed further so as to describe more complex situations than we have previously considered (e.g. section 3). In this section we outline the extended formalism, treating internal dynamics, grammar, and geometry in turn, without however, attempting complete biological generality.

5.1. INTERNAL DYNAMICS

Equations (2) and (3) must be modified so that they apply to each object which has actually been created during development, and so that they apply to state changes arising from cell-cell induction, axon sprouting, or any other processes in the grammar Γ. We indicate how the processes associated with many of the rules in a developmental grammar could be jointly modeled, at a phenomenological level, with continuous- and discrete-time neural nets.

The symbols appearing in the modified equation are indexed as follows. An object existing in the nth generation of a population is indexed by i. The abstract "object index" i is generally interpreted as an ordered sequence of indices $i_1, \ldots, i_n$, which uniquely identifies each object in terms of its lineage. Thus object is parent is associated with the sequence of indices $i_1, \ldots, i_{n-1}$. In the case of mitosis, each index i_k can take on only two values, corresponding to each of the two daughter cells. If these values are represented by 0 and 1, each object has a unique numerical identifier, the binary representation of which describes that object's lineage. For example, suppose that $i = 17$ for some cell. The binary representation of 17 is 10001,

indicating that this cell is a member of the fifth generation. Now suppose that this cell undergoes mitosis. Then the object for the two daughter cells will be $i = 34$ (binary 100010) and $i = 35$ (binary 100011). Even simpler interpretations of i, such as the position co-ordinate of the object, may be useful in special situations. There is also a rule index r, specifying which rule of the grammar is under consideration, and an object type index α which encodes the alternative object types: cell, synapse, fiber compartment, etc. The state vector $\boldsymbol{v}_i$ is indexed by α as well as i: $\boldsymbol{v}_{i_1,\ldots,i_n,\alpha}$. A state vector history is shown in Fig. 1.

In the modified equations we introduce two new dynamical variables A and C. The first of these is an array $A_{i_1,\ldots,i_n,\alpha}(t)$ of 0/1 values. $A_{i_1,\ldots,i_n,\alpha}(t) = 1$ if and only if an object of type α and lineage $i_1, \ldots, i_n$ exists at time t. $C_i^r(t)$ is also a variable taking 0/1 values, which is 1 if and only if object i uses grammar rule r to change its state at time t. The elements of the connection matrices $\mathbf{T}^r$, $\mathbf{T}^r_\alpha$, $\mathbf{T}^r_{\alpha k}$ and $\mathbf{T}^r_{\alpha\beta}$ are parameters to be fit to biological data, and each grammatical rule r has its own matrix of such parameters. Naturally one must ensure that the grammar used to model a biological data set of size d has fewer than d free parameters; therefore sparseness in the $\mathbf{T}$ matrices and in the grammar is favored.

We also introduce a set of sparse 0/1 arrays Γ to describe changes in object type under the action of grammatical rules. In general, a biological object such as a cell or synapse may undergo a variety of transformations. It may be born from a parent or it may die. Between birth and death its internal state will change as a result both of internal dynamics and by interactions with other objects of the same or different type. The grammatical rules described by the set of Γs permit all of these events to be represented mathematically. Birth and death processes are represented by discrete time rules only; changes in internal state and interactions with other objects may be represented by either continuous or discrete time grammar rules. These possibilities taken together amount to six classes of rules. These classes of rules are illustrated in Fig. 3; we now discuss them in detail.

The input and output types of the continuous time rules are described by the arrays $\hat{\Gamma}^r_{\alpha;\alpha}$ and $\hat{\Gamma}^r_{\alpha\beta;\alpha}$; those of the discrete time rules by the arrays $\Gamma^r_{\beta;}$, $\Gamma^r_{\beta;\alpha i_{n+1}}$, $\Gamma^r_{\beta;\alpha}$ and $\Gamma^r_{\beta\gamma;\alpha}$. Each array contains constant 0/1 values describing the input and output types of the rth rule of the grammar; the input and output types are separated by a semicolon in Γs set of subscripts. For example, $\hat{\Gamma}^r_{\alpha;\alpha}$ is zero unless rule r describes an interphase-like process with one input and one output object, both of the same type. For that value of r, $\hat{\Gamma}^r_{\alpha;\alpha} = 1$ and all the other Γ^rs are zero. Similarly, the death of an object is specified by the rule $\Gamma^r_{\beta;}$ which has no output type. Birth may involve the creation of several objects of different types; it is specified by $\Gamma^r_{\beta;\alpha i_{n+1}}$. Here the extra index i_{n+1} following the semicolon is the lineage index for the newly created objects. Events between birth and death may involve either continuous or discrete time rules; these rules may act on one or two objects. Continuous time one-object rules (like interphase) do not alter an object's type α or name i; they are specified by $\hat{\Gamma}^r_{\alpha;\alpha}$. Discrete time one-object rules, in which an object of type β is transformed to a new object of type a, are specified by $\Gamma^r_{\beta;\alpha}$. It often occurs in the combined expression $\Gamma^r_{\beta;\alpha}\delta_{i_{n+1}0}$ which is a special case of $\Gamma^r_{\beta;\alpha i_{n+1}}$ that enforces the requirement that only one object, indexed by $i_{n+1} = 0$, results from

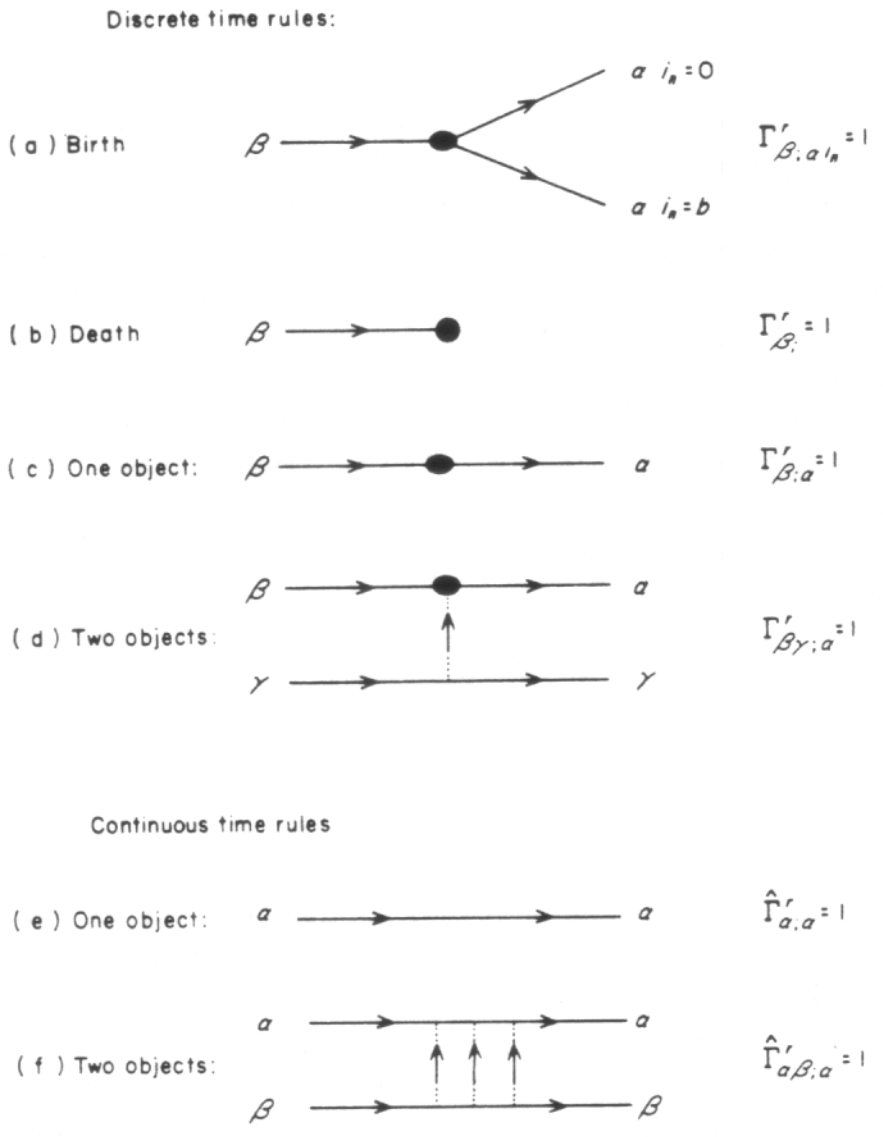

FIG. 3. A diagrammatic representation of the six classes of rules, together with the associated Γ component which is equal to unity if that rule is chosen. In each diagram, the time axis runs in a horizontal direction, and a space axis in the vertical direction. The arrowheads on the solid lines point in the direction of increasing time. The dotted vertical arrows represent a spatial interaction: continuous time rules by three such arrows, discrete time by 1. The dotted arrows point in the direction of the object which has chosen the illustrated rule. A discrete time rule is represented by a filled circle; a continuous time rule by a pair of arrowheads on a solid line without an intervening filled circle. The input and output object types are indicated on the left- and right-hand sides respectively of each diagram. In (a), more than two branches could occur; the branches have lineage indices i_n ranging from 0 to b.

the rule. Note that a type change always implies a change of object index, although for rules not describing birth this change simply requires appending a zero to the object index of the parent.

Similarly there are two types of two-object rules. A rule specified by $\hat{\Gamma}'_{\alpha\beta;\alpha}$ specifies an object of type α interacting with an object or objects of type β with no type change of α. For the interaction of two objects with a type change, we assume for simplicity that only one new object can be created (i.e. i_{n+1} has only one value which we may call 0). Thus, a rule describing a process where an object of type β is induced by an object or objects of type γ to transform into an object of type α is specified by $\Gamma'_{\beta\gamma;\alpha}$. It often occurs in the combined expression $\Gamma'_{\beta\gamma;\alpha}\delta_{i_{n+1}0}$ which

is a special case, limited to the alteration of one object, of an exceedingly general Γ array not required by our grammars: $\Gamma^r_{\beta\gamma;\alpha i_{n+1}}$.

A, C and Γ are responsible for selecting which connection matrix $\mathbf{T}$ and which state vector $\boldsymbol{v}$ are used to produce the new state vector $\boldsymbol{v}_{i,\alpha}$. They do this by multiplying all but the selected $\mathbf{T}$ and $\boldsymbol{v}$ by zero, in a sum over all the possibilities. In an efficient computer implementation of this model, A, C and Γ should be stored as sparse arrays so that terms which are multiplied by zero are never actually computed at all.

To minimize the number of indices appearing in this section's equations, we will uniformly use boldface matrix (e.g. $\mathbf{T}$) and italic/boldface vector (e.g. $\boldsymbol{v}$) notation to suppress the a and b indices that are explicit in previous sections.

Taking into account the above remarks, the required modification of eqn (2) can be written as

$$\boldsymbol{\tau}_\alpha \frac{\mathrm{d}\boldsymbol{v}_{i,\alpha}}{\mathrm{d}t} = A_{i,\alpha}[\boldsymbol{g}(\boldsymbol{u}_{i,\alpha}) - \boldsymbol{\lambda}_\alpha \boldsymbol{v}_{i,\alpha}], \tag{21}$$

where

$$\boldsymbol{u}_{i,\alpha} = \sum_r C^r_i \hat{\Gamma}^r_{\alpha;\alpha} \mathbf{T}^r \cdot \boldsymbol{v}_{i,\alpha} + \sum_r C^r_i \sum_\beta \hat{\Gamma}^r_{\alpha\beta;\alpha} \sum_j \Lambda_{ij} \mathbf{T}^r_\beta \cdot \boldsymbol{v}_{j,\beta} + \boldsymbol{h}_\alpha, \tag{22}$$

and $i = i_1, \ldots, i_n$. That is, the input to the dynamics for $\boldsymbol{v}_{i,\alpha}$ includes an adjustable threshold $\boldsymbol{h}_\alpha$, the current state vector of object i multiplied by a suitable connection matrix describing the interactions of different species v^a within object i, and similar matrix-vector products for each neighboring object j with which i is interacting according to relevant rules of the grammar, modified by a geometric structure factor Λ_{ij} between objects i and j. An example of the specific application of the term for interactions with neighbors can be found in eqn (8), where $\mathbf{T}^{\text{diffusion}}_\beta$ is the identity matrix. Note that if $\boldsymbol{A}_{i,\alpha} = 0$ in eqn (21) (so that object i, if it exists at time t, is not of type α) then $\mathrm{d}\boldsymbol{v}_{i,\alpha}/\mathrm{d}t = 0$; in fact we will ensure that $A_{i,\alpha}(t) = 0 \Rightarrow \boldsymbol{v}_{i,\alpha} = 0$. Consequently we can use the notations $A_i = \sum_\alpha A_{i,\alpha}$ and $\boldsymbol{v}_i = \sum_\alpha \boldsymbol{v}_{i,\alpha}$ for the existence and state vector of object i, regardless of its type.

Discrete time dynamics for mitosis was given in eqn (3); this equation can be generalized to cover any discrete time grammar rule acting at time t to produce new objects an instant later, at time $t + \Delta t$:

$$\boldsymbol{v}_{i_1,\ldots,i_{n+1},\alpha}(t+\Delta t) = A_{i_1,\ldots,i_{n+1},\alpha}(t+\Delta t)[\boldsymbol{v}_{i_1,\ldots,i_n}(t) + \mathbf{R}\boldsymbol{g}(\boldsymbol{u}_{i_1,\ldots,i_{n+1},\alpha}(t))] \tag{23}$$

where at time t

$$\begin{aligned}\boldsymbol{u}_{i_1,\ldots,i_{n+1},\alpha} = &\sum_r C^r_i \sum_\beta \Gamma^r_{\beta;\alpha i_{n+1}} \mathbf{T}^r_{\beta i_{n+1}} \cdot \boldsymbol{v}_{i_1,\ldots,i_n,\beta} \\ &+ \sum_r C^r_i \sum_\beta \Gamma^r_{\beta;\alpha} \delta_{i_{n+1}0} \mathbf{T}^r_\beta \cdot \boldsymbol{v}_{i_1,\ldots,i_n,\beta} \\ &+ \sum_r C^r_i \sum_{\beta\gamma} \Gamma^r_{\beta\gamma;\alpha} \delta_{i_{n+1}0} \sum_j \Lambda_{ij} A_{i_1,\ldots,i_n,\beta} \mathbf{T}^r_{\beta\gamma} \cdot \boldsymbol{v}_{j,\gamma} + \boldsymbol{h}_\alpha,\end{aligned} \tag{24}$$

and again $i = i_1, \ldots, i_n$. Here, the input to the dynamics for $\boldsymbol{v}_{i,\alpha}$ includes the adjustable threshold $\boldsymbol{h}_\alpha$, as well as three additional terms. The first term is a birth rule involving the state vector of object is parent, which was of type β, multiplied

by the relevant connection matrix as determined by the choice of discrete time grammar rule r. Note that the connection matrix in this term is indexed by both β and i_{n+1}. This is necessary because each object born by this rule will in general have its own connection matrix. The second term is a one-object rule by which object i changes from type β to type α. Note that the connection matrix in this term is indexed by β only, and not by i_{n+1}, since only one new object is involved. The third term is a two-object rule in which an object of type β is induced to transform into an object of type α by interacting with one or more objects of types γ. The state vector of the inducing objects j of types γ are each multiplied by the appropriate connection matrix $\mathbf{T}^r_{\beta\gamma}$, which must be selected not only by the rule r and inducing type γ, but also by the existence variable $A_{i_1,\ldots,i_n,\beta}$, which picks out the appropriate parental object type for the connection matrix. The death rule $\Gamma^r_{\beta;}$ is not included in eqn (24), since it exerts its effects solely through the dynamics of $A_{i,\alpha}$. These equations do not yet determine $A_{i,\alpha}$, C^r_i, or Λ_{ij}, whose dynamics we consider next.

5.2. GRAMMAR

The set of objects existing in the nth generation at time t is specified by the set of $A_{i_1,\ldots,i_n,\alpha}(t)$ for all possible lineages. In conjunction with the grammar Γ and the choice $C^r_{i_1,\ldots,i_n}(t)$ of grammar rules to be applied at time t to an object of generation n, $A_{i_1,\ldots,i_n,\alpha}(t)$ determines which objects exist subsequently. The dynamics of A is given by:

$$\begin{aligned} A_{i_1,\ldots,i_{n+1},\alpha}(t+\Delta t) = &\sum_r C^r_{i_1,\ldots,i_{n+1}}(t)\hat{\Gamma}^r_{\alpha;\alpha}A_{i_1,\ldots,i_{n+1},\alpha}(t) \\ &+\sum_r C^r_{i_1,\ldots,i_n}(t)\sum_\beta \Gamma^r_{\beta;\alpha i_{n+1}}A_{i_1,\ldots,i_n,\beta}(t) \\ &+\Theta\left[\sum_r C^r_{i_1,\ldots,i_n}(t)\left(\sum_\beta \Gamma^r_{\beta;\alpha}+\sum_{\beta\gamma}\Gamma^r_{\beta\gamma;\alpha}\right)\delta_{i_{n+1}0}A_{i_1,\ldots,i_n,\beta}(t)\right], \end{aligned} \tag{25}$$

where the function $\Theta(n)=0$ for $n\le 0$, $\Theta(n)=1$ for $n>0$. In this dynamical equation, the first summand represents a one-object continuous time rule where no new objects are created. The second term represents a birth rule, while the last two summands represent a change of object type respectively by one- and two-object discrete time rules. Continuous time two-object rules do not figure in this equation because they have no effect on the number of interacting objects. Death rules are not explicitly listed either, but choosing a death rule will force the right side of the equation to vanish by virtue of the constraint eqn (26), introduced below. The use of Θ insures that A will only take on the values 0 or 1.

We place two constraints on A. At a given time, an object can belong to only one generation, so $A_{i_1,\ldots,i_n}(t)=1$ implies that $A_{i_1,\ldots,i_m}(t)=0$ for $m<n$. If a parent biological object (indexed by $i_1,\ldots,i_{n-1}$) actually survives the creation of its progeny, we simply relabel it as one of the progeny (conventionally with index $i_n=0$) and consider the old object to be non-existent, so that $A_{i_1,\ldots,i_{n-1}}=0$. The constraint that a given object be of only one type at a time is expressed by $A_i=\sum_\alpha A_{i,\alpha}\in\{0,1\}$. In

the Appendix we show that these properties are preserved by eqn (25). With respect to $\boldsymbol{v}$, note that for fixed i, $\boldsymbol{v}_{i,\alpha}$ is non-zero for only one value of α, so that $\boldsymbol{v}_i = \sum_\alpha \boldsymbol{v}_{i,\alpha}$. Since $A_{i,\alpha}$ multiplies $\boldsymbol{v}_{i,\alpha}$ in eqns (21) and (23), this fact is preserved by the dynamics of $\boldsymbol{v}$.

At each time, each object i is subject to one or several grammar rules; this choice is encoded by the C_i^r variables. We place certain constraints upon which combinations of rules can be selected at a given time t. For simplicity, we allow only one birth or death rule to operate at a time. An object typically changes state according to its own internal dynamics and through interactions with a number of neighbors. Hence, we allow any finite number of continuous time two-object rules to operate simultaneously, but only one continuous time one-object rule. Similarly, any finite number of discrete time rules with one output object may operate simultaneously, but they must all have the same output type α. This condition places a constraint on the dynamics of C in terms of A, expressed by requiring

$$A_{i,\beta} = \sum_{\substack{r=\text{birth rules} \\ \text{from } \beta}} C_i^r + \sum_\alpha \Theta\left[\sum_{\substack{r=\text{discrete time one- and} \\ \text{two-object rules from } \beta \to \alpha}} C_i^r \right] + \sum_{\substack{r=\text{death rules} \\ \text{from } \beta}} C_i^r + \sum_{\substack{r=\text{continuous time,} \\ \text{one-object rules from } \beta}} C_i^r . \tag{26}$$

The use of the function Θ, introduced in eqn (25), expresses the idea that any non-zero number of discrete time rules can be chosen, as long as they all produce the same type of object. As with eqn (25), two-object continuous time rules are not enumerated in (26) because they do not affect the number of objects. Equation (42) in the Appendix is a completely explicit formulation of (26) above.

The dynamics of C^r can be summarized by its effect, which is that each C^r has an associated strength S^r, computed locally at object i, and the decision as to which rule to invoke goes to the set of variables with the greatest summed strength. This procedure can be expressed as the optimization of an objective function: For all existing objects [at a given time t and for all i for which $A_i(t) \neq 0$], optimize

$$E_i[C] = \sum_r C_i^r S_i^r , \tag{27}$$

subject to the constraint (26).

The entire model would be complete if we knew the rule strengths S_i^r. Each object i of type α has for each value of r a rule strength

$$S_i^r = \sum_\alpha S_{i,\alpha}^r A_{i,\alpha} , \tag{28}$$

where

$$S_{i,\alpha}^r = \boldsymbol{v}_{i,\alpha} \cdot \boldsymbol{s}_\alpha^r + \theta_\alpha^r + \sum_{j\beta} \Lambda_{ij} (\boldsymbol{v}_{i,\alpha} \cdot \boldsymbol{s}_{\alpha\beta}^r \cdot \boldsymbol{v}_{j,\beta} + \boldsymbol{s}_\beta^r \cdot \boldsymbol{v}_{j,\beta}) . \tag{29}$$

The connection between the rule strength S_i^r for object i and that object's state vector $\boldsymbol{v}_i$ is encoded in the strength connection $\boldsymbol{s}_\alpha^r$, while θ_α^r is that part of the rule strength not dependent on the state variables. For example, suppose $\boldsymbol{s}_\alpha^r$ is a unit vector with just one non-zero component, and θ_α^r and the other $\boldsymbol{s}$s are zero. In this

case S_i^r is the size of the corresponding component of $\boldsymbol{v}_i$, so that the concentration of just one protein within object i determines the strength of a rule in the grammar at that object. θ_α^r often occurs as a threshold; two biological situations where this is the case are discussed in section 6. The part of the rule strength S_i^r for object i depending on interactions with other objects j is given by the sum over j and β in eqn (29). In this sum there is a contribution from the interaction of the state vectors of objects i and j controlled by the strength connection matrix $\boldsymbol{s}_{\alpha\beta}^r$, as well as a contribution directly from the state vector of object j, mediated by $\boldsymbol{s}_\beta^r$. All the contributions to S_i^r from other objects j are modified by a geometric structure factor Λ_{ij}.

The parameters $\boldsymbol{s}_\alpha^r$, θ_α^r, $\boldsymbol{s}_{\alpha\beta}^r$, and $\boldsymbol{s}_\beta^r$ are analogous to the connection matrices $\mathbf{T}$ and the thresholds $\boldsymbol{h}_\alpha$ in that they are to be learned by fitting to biological data. The term of eqn (29) containing $\boldsymbol{s}_{\alpha\beta}^r$ is motivated by a quadratic mass-action model of a number of reactions between chemical species from objects i and j. The terms containing the strength connections $\boldsymbol{s}_\alpha^r$ and $\boldsymbol{s}_\beta^r$ may be thought of as reactions between chemical species specifying the state of object i or object j respectively and other chemical species in the respective objects which do not participate in the regulatory circuit and whose concentrations remain constant.

5.3. GEOMETRY

The dynamical interactions of the fundamental objects that constitute a developing organism will be influenced by the arrangement of these objects in space; likewise the dynamics will influence the spatial organization. To treat this problem, we consider the coupling of spatial geometry to the internal dynamics and grammatical sectors of the model. Geometry enters the model only in eqns (24) and (29), through the geometric structure factor Λ_{ij} between two objects i and j. It is necessary to include the dynamics of Λ_{ij} in any application, or the model will be incomplete and incapable of being simulated or fit to biological data. Our phenomenological modeling framework would be best advanced by a generic, approximate, and heavily parameterized model of the dynamics of Λ_{ij}, such as the model just described for internal state changes and object birth and death according to the rules of a grammar. As yet, we have not developed such a model.

This problem was circumvented in the blastoderm model of section 3 by using an application-specific, non-learnable geometric factor for Λ_{ij} that results from the physics of a discrete diffusion operator and a very simple geometry of cell division. The fact that geometry influences the rest of the model through Λ_{ij} encourages such a simple strategy in biological situations where the relevant geometry is already well-understood and modeled. Otherwise, one needs the missing generic, phenomenological, and tunable equations for Λ_{ij}. In this section we present a preliminary sketch of dynamical equations for Λ_{ij} which is, unfortunately, too simplified to model actual biological development.

The idea of this model is that Λ_{ij} is directly determined as a simple function of the relative positions and a small set of shape variables $\boldsymbol{\rho}_i$ and $\boldsymbol{\rho}_j$ (such as centroid, orientation and aspect ratio) for objects i and j. These position and shape variables

are governed by a differential equation involving mechanical forces, which are determined by Λ_{ij} and by the internal state variables $v_{i,\alpha}$ of the relevant objects. A discrete time grammar rule must also determine the new geometric and shape variables in terms of the old ones.

Since $\sum_\alpha A_{i,\alpha} \in \{0, 1\}$, we can define the geometric structure factor between objects i and j as $\Lambda_{ij} = \sum_{\alpha\beta} A_{i,\alpha} A_{j,\beta} \Lambda_{i\alpha,j\beta}$ (only one summand will be non-zero). A simple choice would be to assume that it is a function of the relative positions x_i and shape variables ρ_i of objects i and j, such as $\Lambda_{i\alpha,j\beta} = \Lambda_{\alpha\beta}(x_i - x_j, \rho_i, \rho_j)$. (This also allows the Λ_{ij} dynamics to depend on object type.) To prevent the number of shape parameters from growing large, we assume that extended objects such as axons are modeled as interacting collections of elementary shapes such as cable segments. An even simpler choice would be to assume that Λ_{ij} is a function of the relative positions alone, $\Lambda_{i\alpha,j\beta} = \Lambda_{\alpha\beta}(x_i - x_j)$. Then the model would again be complete if we knew the geometry of the objects as specified by their positions, x_i.

Spatial information can be incorporated into the model by specifying an equation of motion for the center of mass x_i of each object. For the present we assume a "slow-motion" or "viscous-dominated" approximate dynamics:

$$\frac{dx_i}{dt} = \eta F_i, \tag{30}$$

with

$$F_i = F_i^{\text{contact}} + F_i^{\text{other}}. \tag{31}$$

We think of the two indicated contributions to the force as follows: F_i^{contact} results from gradients of adhesion, morphogens, and in general any state-vector-dependent force between two objects, and F_i^{other} summarizes other forces not discussed here. We also impose detailed force balance: $\sum_i F_i^{\text{any}} = 0$. Consequently any background or substrate objects will have to be explicitly included in the model.

A state-vector-dependent force between two objects, such as an adhesion or morphogen gradient, is determined by a quadratic form involving the two state vectors:

$$F_i^{\text{contact}} = \nabla_{x_i} a_i(x_i),$$

where

$$a_i(x_i) = \sum_j a_{ij} \Lambda_{\alpha(i),\beta(j)}(x_i - x_j), \tag{32}$$

and

$$a_{ij} = \sum_{\alpha\beta} v_{i,\alpha} \cdot Q_{\alpha\beta} \cdot v_{j,\beta}.$$

Here $a_i(x_i)$ is a potential energy, and its gradient is a mechanical force. Note that Λ_{ij} symmetrizes a_{ij} and therefore detailed force balance is satisfied for F_i^{contact}. The parameters $Q_{\alpha\beta}$ are to be trained on biological data.

Once again we caution that these equations are only an illustration of the possibility of a generic, phenomenological and learnable model for the geometric structure factors, not an example of one. For example the much more complete

cell-cell adhesion models reviewed in Bell (1988) would not easily be approximated by eqn (32).

6. Notes on the Use of the Formalism

In this section we discuss a few mechanisms which are known to be important in biological development. We show how these interactions can be expressed in the modeling framework outlined in section 5.

Recent advances in the understanding of cell cycle control in *Drosophila* (O'Farrell *et al.*, 1989) indicate that the decision to enter the mitotic cycle is governed by the level of the product of the *string* gene, such that whenever the concentration of *string* product is above a critical threshold, the cell enters mitosis. The level of *string* product is itself under the control of pattern formation genes (O'Farrell *et al.*, 1989). This situation can be modeled as follows. Consider a single nucleus (i.e. fix i and α). In eqn (29), set the strength connections s for mitosis to be

$$s_{\alpha=\text{nucleus}}^{a,r=\text{mitosis}} = \begin{cases} 1 & \text{iff } a = \textit{string} \\ 0 & \text{otherwise.} \end{cases} \tag{33}$$

For interphase,

$$s_{\alpha=\text{nucleus}}^{a,r=\text{interphase}} = 0 \quad \text{for all } a. \tag{34}$$

Then eqn (29) becomes

$$S_i^{\text{mitosis}} = v_{i,\text{nucleus}}^{\textit{string}} \tag{35}$$

$$S_i^{\text{interphase}} = \theta_{\text{nucleus}}^{\text{interphase}}, \tag{36}$$

where $\theta_{\text{nucleus}}^{\text{interphase}}$ is the threshold above which *string* initiates the mitotic process, and $\theta_{\text{nucleus}}^{\text{mitosis}} = 0$. We maximize eqn (27), which now reduces to

$$E[\text{mitosis and interphase}] = \sum_{r=1}^{2} C_i^r S_i^r. \tag{37}$$

Since we already assumed the existence of nucleus i, constraint (25) is satisfied. Because C^{mitosis} is a birth rule, constraint (26) dictates that only one of the two Cs in eqn (37) can be equal to 1. Hence, maximizing E reduces to picking out which quantity is larger: $v^{\textit{string}}$ or θ^{mitosis}. In other words nucleus i will divide when the level of *string* product is greater than the threshold for initiating mitosis. This provides a way of extending the model of the *Drosophila* blastoderm presented in section 3 to later times when mitosis is under the autonomous control of the zygotic genome.

As a second example, we consider cell-cell induction. Induction is a process where one cell changes type as a result of receiving a signal from a neighboring cell. Recent work has begun to unravel this process in terms of specific signaling factors. For example, Smith (1987) has described a factor secreted by a *Xenopus* cell line that induces animal pole cells from a *Xenopus* embryo to form mesoderm. This factor, apparently a small protein, may well be the agent by which cells of the vegetal pole induce mesoderm in normal development. Here our intention is not to

attempt to describe all of the dynamic players involved in the induction process, but rather to sketch out how the signaling step of induction might be described by the formalism presented in section 5. For brevity we drop the object index i in the following; since we do not need to keep track of distinct objects of the same type there is no loss of essential information.

We have a target cell of type "animal pole", or $\alpha = \text{ap}$, which is to be induced to become a cell of type mesoderm, or $\alpha = \text{me}$. There is also an inducing cell of type XTC (the tissue culture line used by Smith). We assume Smith's factor to be equivalent to the native factor, and denote the XTC cell type by $\alpha = \text{vp}$, for vegetal pole. One element of $\boldsymbol{v}_{\text{vp}}$ is the concentration of the inducing factor $v_{\text{vp}}^{\text{inducer}}$; an element of $v_{\text{ap}} = v_{\text{ap}}^{\text{receptor}}$ is the concentration of a putative receptor for the inducing agent. Smith directly observed one element of v_{me}, namely the level of myosin expression (a marker for mesoderm) $v_{\text{me}}^{\text{myo}}$.

We wish to model induction via a two-object discrete time rule. In the present sketch, we consider only two choices, "induction" and "stasis". The former is a discrete time rule and the latter a continuous time rule. The induction of a type change in one cell by its neighbor can be expressed directly by including a cell-cell induction rule in the grammer Γ. The type change encoded in this rule is specified by setting $\Gamma^{r}_{\beta\gamma;\alpha} = \Gamma^{\text{induction}}_{\text{vpap;me}} = 1$. We assume the rule strength for stasis to be fixed, so that eqn (29) becomes, for this rule,

$$S^{\text{stasis}} = \theta^{\text{stasis}}_{\text{ap}}. \tag{38}$$

The decision for induction presumably depends on both the inducer and its receptor, as well as a measure of how tightly the two couple. Comparison with eqn (29) indicates that

$$S^{\text{induction}} = \Lambda s^{\text{induction}} v^{\text{inducer}} v^{\text{receptor}}, \tag{39}$$

where all indices on s other than the rule index have been suppressed. The arguments given with regard to optimizing $E[C]$ in eqn (37) carry over to this rule choice almost unchanged. Since we are considering a discrete time and a continuous time rule, only one can be selected, and the rule choice amounts to making a decision about a threshold. Hence, it follows that induction will occur whenever the product $\Lambda s^{\text{induction}} v^{\text{inducer}} v^{\text{receptor}}$ is greater than the threshold θ^{stasis}.

The dynamics of the induction rule are specified in eqn (24) by the connection matrix $\mathbf{T}^{\text{induction}}_{\text{vpap}}$. This connection is a direct input from the inducing cell's state vector to the dynamics of the target cell's state vector. The only non-zero element of this matrix clearly implied by the work of Smith (1987) is that between $v_{\text{vp}}^{\text{inducer}}$ and $v_{\text{me}}^{\text{myosin}}$; the latter shows an increasing sigmoidal response to increasing doses of inducer. Such a coupling can be modeled using the thresholding function g, but detailed comparisons of our model with the dynamics of this system must await further data.

7. Discussion

In this paper we have presented a framework describing the dynamics of development. Advances in biochemical technology have made it possible to measure many

of the state variables that determine development. As the methods for obtaining information about these biological regulatory substances improve, it will become increasingly important to understand how these substances, acting in concert, control growth and development.

We think it possible that the modeling framework presented here will allow these questions to be addressed in a phenomenological but precise and systematic manner. Modern biology is very much a science of contingencies. Many experiments involve a perturbation of the concentration of some biochemical regulator, followed by a study of the consequences of that perturbation in terms of the consequent changes in concentrations of other regulators, effects on growth, or effects on morphology. The framework presented here is designed to discover and express correlation in data, even when knowledge of underlying mechanisms is incomplete.

We have presented a number of examples of how this formalism can be used to model important developmental systems and events. These include gene regulation in the *Drosophila* blastoderm, mitotic control following gastrulation in the same animal, mesoderm induction in *Xenopus*, and programmed cell death. The formalism is intended to be used to generate specific models, the parameters of which are found by means of a non-linear least squares fit to time series data. The details of the procedure will be presented elsewhere, but preliminary numerical results on the *Drosophila* blastoderm (Reinitz *et al.*, 1991) indicate that our approach is feasible and has predictive value.

The modeling framework contains three sectors. Two of these, concerned with internal dynamics and grammar, may well be generic. This question can only be settled by comparison to data in as many systems as possible. The third sector, that of geometry, is at yet incomplete, although usable for systems whose geometry is relatively simple. A full treatment of the geometrical problem is an important topic for future work.

The dynamical equations we employ, while phenomenological, can be derived from simple but explicit biochemical mechanisms. The understanding of specific biochemical reaction *mechanisms*, particularly in eucaryotes, has lagged behind the technology for assaying the *levels* of the reacting species. For this reason we think that that discussion of biochemical mechanism given in section 4 is as faithful to the underlying mechanisms as the data will permit in the near future.

We thank M. Levine and S. Letovsky for encouragement and stimulating conversations, and E. Sklar, P. Hackett, D. Hackett and C. Garrett for essential logistical support.

REFERENCES

AKAM, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**, 1-22.

BAUM, E. & HAUSSLER, D. (1989). What size net gives valid generalization? In: *Neural Information Processing Systems* Vol. 1, p. 81. Palo Alto, CA: Morgan Kaufmann.

BELL, G. I. (1988). Models of cell adhesion involving specific binding. In: *Physical Basis of Cell-Cell Adhesion* (Bongrand, P., ed.) chapter 10. Boca Raton, FL: CRC Press.

CARROLL, S. B., LAUGHON, A. & THALLEY, B. (1988). Expression, function, and regulation of the *hairy* segmentation protein in the *Drosophila* embryo. *Genes and Development* **2**, 883-890.

CARROLL, S. B. & SCOTT, M. P. (1986). Zygotically active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. *Cell* **45,** 113-126.

EDGAR, B. A., ODELL, G. M. & SCHUBIGER, G. (1989). A genetic switch, based on negative regulation, sharpens stripes in *Drosophila* embryos. *Dev. Gen.* **10,** 124-142.

FOE, V. A. & ALBERTS, B. M. (1983). Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis. *J. Cell Sci.* **61,** 31-70.

FRASCH, M. & LEVINE, M. (1987). Complementary patterns of *even-skipped* and *fushi tarazu* expression involve their differential regulation by a common set of segmentation genes in *Drosophila. Genes Dev.* **1,** 981-995.

GOODWIN, B. & KAUFFMAN, S. (1990). Spatial harmonics and pattern specification in early *Drosophila* development. Bifurcation sequences and gene expression. *J. theor. Biol.* **144,** 303-319.

HILL, T. L. (1985). *Cooperativity Theory in Biochemistry: Steady-State and Equilibrium Systems.* Berlin: Springer-Verlag.

HOPFIELD, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Nat. Acad. Sci. U.S.A.* **81,** 3088-3092.

INGHAM, P. W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila. Nature, Lond.* **335,** 25-34.

JACKLE, H., TAUTZ, D., SCHUH, R., SEIFERT, E. & LEHMANN, R. (1986). Cross-regulatory interactions among the gap genes of *Drosophila. Nature, Lond.* **324,** 668-670.

LACALLI, T. C., WILKINSON, D. A. & HARRISON, L. G. (1988). Theoretical aspects of stripe formation in relation to *Drosophila* segmentation. *Development* **104,** 105-113.

LINDENMAYER, A. (1968). Mathematical models for cellular interaction in development, parts I and II. *J. theor. Biol.* **18,** 280-315.

MEINHARDT, H. (1986). Hierarchical inductions of cell states: a model for segmentation in *Drosophila. J. Cell Sci.* (Suppl.) **4,** 357-381.

MJOLSNESS, E., SHARP, D. H. & ALPERT, B. K. (1989). Scaling, machine learning, and genetic neural nets. *Advan. appl. Math.* **10,** 137-163.

MULLER, H. J. (1932). Further studies on the nature and causes of gene mutations. In: *Sixth International Congress of Genetics* Vol. 1, p. 213-255.

O'FARRELL, P. H., EDGAR, B. A., LAKICH, D. & LEHNER, C. F. (1989). Directing cell division during development. *Science* **246,** 635-640.

PANKRATZ, M. J., HOCH, M., SEIFERT, E. & JACKLE, H. (1989). *Kruppel* requirement for *knirps* enhancement reflects overlapping gap gene activities in the *Drosophila* embryo. *Nature, Lond.* **341,** 337-340.

PEARMUTTER, B. A. (1989). Learning state space trajectories in recurrent neural networks. *Neural Computation,* **1**(2), 263-269.

REINITZ, J. & LEVINE, M. (1990). Control of the initiation of homeotic gene expression by the gap genes *giant* and *tailless* in *Drosophila. Dev. Biol.* **140,** 57-72.

REINITZ, J., MJOLSNESS, E. & SHARP, D. H. (1991). A connectionist model of the *Drosophila* blastoderm. In: *The Organization of Organisms: Proceedings of a Workshop at the Santa Fe Institute* (Baskin, A. & Mittenthal, J., eds) Redwood City, CA: Addison-Wesley.

RUMMELHART, D. E., HINTON, G. E. & WILLIAMS, R. J. (1986). Learning internal representations by error propagation. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Cambridge, MA: MIT Press.

RUSHLOW, C., HARDING, K. & LEVINE, M. (1987). Hierarchical interactions among pattern-forming genes in *Drosophila.* In: *Banbury Report* 26: *Development Toxicity: Mechanisms and Risk.* Cold Spring Harbor: Cold Spring Harbor Press.

SMITH, J. C. (1987). A mesoderm-inducing factor is produced by a *Xenopus* cell line. *Development* **99,** 3-14.

APPENDIX

Properties of the Grammatical Rules

In this appendix, we prove that certain properties of A are preserved in time under the action of the grammatical rules. To do this we need to state explicitly certain properties of the grammar that were described in words in section 5. In that section, and in Fig. 3, we describe six classes of grammatical rules. Every rule

belongs to only one class, and each rule has unique input and output types, with the exception that the birth class of rule may have several output types. So for all r it is true that

$$1 = \sum_{\alpha\beta} \Gamma^r_{\beta;\alpha} + \sum_{\beta} \Gamma^r_{\beta;} + \sum_{\beta} \Theta\left(\sum_{\alpha} \sum_{i_n} \Gamma^r_{\beta;\alpha i_n} \right) + \sum_{\alpha\beta\gamma} \Gamma^r_{\beta\gamma;\alpha} + \sum_{\alpha} \hat{\Gamma}^r_{\alpha;\alpha} + \sum_{\alpha\beta} \hat{\Gamma}^r_{\alpha\beta;\alpha}. \quad (40)$$

In the case of the birth rule, each of the outputs has a unique output type, so that for all r, β, and i_n

$$\sum_{\alpha} \Gamma^r_{\beta;\alpha i_n} \leq 1. \quad (41)$$

Equation (26), written out in full, becomes

$$A_{i_1,\ldots,i_n,\beta} = \sum_r C^r_{i_1,\ldots,i_n} \Theta\left[\sum_{\alpha, i_{n+1}} \Gamma^r_{\beta;\alpha i_{n+1}} \right] + \sum_{\alpha} \Theta\left[\sum_r C^r_{i_1,\ldots,i_n} \Theta\left(\Gamma^r_{\beta;\alpha} + \sum_{\gamma} \Gamma^r_{\beta\gamma;\alpha} \right) \right]$$
$$+ \sum_r C^r_{i_1,\ldots,i_n} \Gamma^r_{\beta;} + \sum_r C^r_{i_1,\ldots,i_n} \sum_{\alpha} \hat{\Gamma}^r_{\alpha;\alpha}. \quad (42)$$

Equation (25) gives the dynamics of $A_{i,\alpha}$. We will also need an explicit dynamical equation for $A_i = \sum_\alpha A_{i,\alpha}$. It is

$$A_{i_1,\ldots,i_{n+1}}(t+\Delta t) = \sum_r C^r_{i_1,\ldots,i_{n+1}}(t) \sum_{\alpha} \hat{\Gamma}^r_{\alpha;\alpha} A_{i_1,\ldots,i_{n+1},\alpha}(t)$$
$$+ \sum_r C^r_{i_1,\ldots,i_n}(t) \sum_{\alpha\beta} \Gamma^r_{\beta;\alpha i_{n+1}} A_{i_1,\ldots,i_n,\beta}(t)$$
$$+ \sum_{\alpha} \Theta\left[\sum_r C^r_{i_1,\ldots,i_n}(t) \left(\sum_{\beta} \Gamma^r_{\beta;\alpha} + \sum_{\beta\gamma} \Gamma^r_{\beta\gamma;\alpha} \right) \delta_{i_{n+1}0} A_{i_1,\ldots,i_n,\beta}(t) \right] \quad (43)$$

With these equations we can prove that the grammatical rules preserve important properties of A. The proof holds for discrete time dynamics with $t_p = p\Delta t$, with p an integer.

Theorem 1: For every ordered sequence $i_1, \ldots, i_n$, and for all α, the following three statements are true at all times $t = p\Delta t$: (a) $A_{i_1,\ldots,i_n,\alpha}(t) \in \{0, 1\}$. (b) If $A_{i_1,\ldots,i_n}(t) = 1$, then $A_{i_1,\ldots,i_m}(t) = 0$ for all $m < n$. (c) $A_{i_1,\ldots,i_n}(t) \in \{0, 1\}$.

Propositions (a)-(c) of Theorem 1 are true by definition at $t = 0$; we show that if the propositions are true at time t, they are true at $t + \Delta t$. Thus the proof is by induction on the discrete variable t (or p).

The proof of proposition (a) reduces to that of (c), since we show that if (c) is true at $t + \Delta t$, then (a) is also true at $t + \Delta t$. To see this, note that by definition of $A_{i_1,\ldots,i_n}(t)$,

$$1 \geq A_{i_1,\ldots,i_n}(t+\Delta t) = \sum_{\alpha} A_{i_1,\ldots,i_n,\alpha}(t+\Delta t). \quad (44)$$

At time $t = 0$, all the As are 0 or 1, and eqn (25) has only non-negative terms, so $A_{i,\alpha}$ can never be negative. If a sum over non-negative integral terms is less than or equal to 1, no term can be greater than 1. Hence, $A_{i,\alpha}(t+\Delta t) \in \{0, 1\}$.

Next consider proposition (b). There are two possible cases. First, suppose that $A_{i_1,\ldots,i_n}(t)=1$. By hypothesis, $A_{i_1,\ldots,i_m}(t)=0$ for any $m<n$. Then by eqn (43) with $n+1=m$, $A_{i_1,\ldots,i_m}(t+\Delta t)=0$ for all $m<n$.

Now suppose that $A_{i_1,\ldots,i_n}(t)=0$, and also assume that $A_{i_1,\ldots,i_n}(t+\Delta t)=1$. Then by (43), we find $A_{i_1,\ldots,i_{n-1}}(t)=1$, so by induction $A_{i_1,\ldots,i_m}(t)=0$ for $m<n-1$. Then by (43) we now find that $A_{i_1,\ldots,i_m}(t+\Delta t)=0$ for all $m<n-1$. We need now only consider the case of $m=n-1$.

Assume the contrary, that $A_{i_1,\ldots,i_{n-1}}(t+\Delta t)=1$. We showed above that $A_{i_1,\ldots,i_{n-2}}(t)=0$. This, in conjunction with (43), implies that $\sum_r C_{i_1,\ldots,i_{n-1}}(t)\hat{\Gamma}^r_{\alpha;\alpha}A_{i_1,\ldots,i_{n-1},\alpha}(t)=1$ for some α, and hence there is a continuous time (here applied to a set of discrete time steps) single-object rule r such that $C^r_{i_1,\ldots,i_{n-1}}(t)=1$. By (42), $C^r_{i_1,\ldots,i_{n-1}}(t)=0$ for all discrete time rules r. Consider (43) with $n-1$ substituted for n. We are considering the case where $A_{i_1,\ldots,i_n,\alpha}(t)=0$ for all α, so the first term of (43) vanishes. But we just showed that for all discrete time r, $C^r_{i_1,\ldots,i_{n-1}}(t)=0$, and so the second and third terms of (43) vanish also. Hence (43) reduces to $A_{i_1,\ldots,i_{n-1},\alpha}(t+\Delta t)=0$ for all α, and so $A_{i_1,\ldots,i_{n-1}}(t+\Delta t)=0$, contradicting the assumption.

Lastly, we prove (c). At time t, $1\geq A_{i_1,\ldots,i_n}(t)=\sum_\beta A_{i_1,\ldots,i_n,\beta}(t)$ by the inductive hypothesis. $A_{i_1,\ldots,i_n,\beta}(t)$ is simply the left side of eqn (42), and every term of (42) is non-negative. We sum the right-hand side of (42) over β and note that any subset of its summands must be less than or equal to the entire sum. Hence

$$1\geq\sum_{r,\beta}C^r_{i_1,\ldots,i_n}(t)\Theta\left[\sum_{\alpha,i_{n+1}}\Gamma^r_{\beta;\alpha i_{n+1}}\right]+\sum_{\alpha\beta}\Theta\left[\sum_r C^r_{i_1,\ldots,i_n}(t)\Theta\left[\Gamma^r_{\beta;\alpha}+\sum_\gamma\Gamma^r_{\beta\gamma;\alpha}\right]\right]. \tag{45}$$

The two Θs in the second term are redundant, so we may write

$$1\geq\sum_{r,\beta}C^r_{i_1,\ldots,i_n}(t)\Theta\left[\sum_{\alpha,i_{n+1}}\Gamma^r_{\beta;\alpha i_{n+1}}\right]+\sum_{\alpha}\Theta\left[\sum_r C^r_{i_1,\ldots,i_n}(t)\left[\sum_\beta\Gamma^r_{\beta;\alpha}+\sum_{\beta\gamma}\Gamma^r_{\beta\gamma;\alpha}\right]\right]. \tag{46}$$

In the first term, we may replace the sum over i_{n+1} by any of its elements and preserve the inequality, since all terms are non-negative. Hence, for any i_{n+1},

$$1\geq\sum_{r,\beta}C^r_{i_1,\ldots,i_n}(t)\Theta\left[\sum_\alpha\Gamma^r_{\beta;\alpha i_{n+1}}\right]+\sum_{\alpha}\Theta\left[\sum_r C^r_{i_1,\ldots,i_n}(t)\left[\sum_\beta\Gamma^r_{\beta;\alpha}+\sum_{\beta\gamma}\Gamma^r_{\beta\gamma;\alpha}\right]\right]. \tag{47}$$

The argument of Θ in the left-hand term is now the left-hand side of eqn (41). This itself is never greater than unity, we may drop the Θ. We now have, for all i_{n+1},

$$1\geq\sum_r C^r_{i_1,\ldots,i_n}(t)\sum_{\alpha\beta}\Gamma^r_{\beta;\alpha i_{n+1}}+\sum_{\alpha}\Theta\left[\sum_r C^r_{i_1,\ldots,i_n}(t)\left[\sum_\beta\Gamma^r_{\beta;\alpha}+\sum_{\beta\gamma}\Gamma^r_{\beta\gamma;\alpha}\right]\right]. \tag{48}$$

We can multiply any term of the right-hand side of the above equation by any variable that takes on only the values of 0 or 1 and still maintain the inequality. Hence, we find that

$$1\geq\sum_r C^r_{i_1,\ldots,i_n}(t)\sum_{\alpha\beta}\Gamma^r_{\beta;\alpha i_{n+1}}A_{i_1,\ldots,i_n,\beta}(t)$$

$$+\sum_{\alpha}\Theta\left[\sum_r C^r_{i_1,\ldots,i_n}(t)\left[\sum_\beta\Gamma^r_{\beta;\alpha}+\sum_{\beta\gamma}\Gamma^r_{\beta\gamma;\alpha}\right]\right]\delta_{i_{n+1}0}A_{i_1,\ldots,i_n,\beta}(t).$$

Note that this is in fact the last two terms of eqn (43). Look now at the first term of (43). Consider the sum over α as an inner product between Γ and A. All of the components of A, C and Γ are non-negative, so that

$$\sum_{\alpha,\alpha'} \delta_{\alpha\alpha'} \hat{\Gamma}^r_{\alpha;\alpha} A_{i_1,\ldots,i_{n+1},\alpha'}(t) \leq \sum_{\alpha,\alpha'} \hat{\Gamma}^r_{\alpha;\alpha} A_{i_1,\ldots,i_{n+1},\alpha'}(t) \tag{50}$$

whence

$$\sum_r C_{i_1,\ldots,i_{n+1}}(t) \sum_\alpha \hat{\Gamma}^r_{\alpha;\alpha} A_{i_1,\ldots,i_{n+1},\alpha}(t) \leq \left(\sum_r C^r_{i_1,\ldots,i_{n+1}} \sum_\alpha \hat{\Gamma}^r_{\alpha;\alpha} \right) \left[\sum_{\alpha'} A_{i_1,\ldots,i_{n+1},\alpha'}(t) \right]. \tag{51}$$

On the right-hand side of this inequality, each of the two factors can be no larger than unity. The left-hand factor appears as one of the terms in (42), as one of several non-negative integral summands whose sum is at most 1. The right-hand factor cannot be greater than one by the inductive hypothesis.

We have now shown that in eqn (43), the first term on the right-hand side can be equal to at most 1, and that the second and third terms have a total value of at most 1. We now show that if the first term is 1, the second and third terms must vanish. Suppose that

$$\sum_r C_{i_1,\ldots,i_{n+1}}(t) \sum_\alpha \hat{\Gamma}^r_{\alpha;\alpha} A_{i_1,\ldots,i_{n+1},\alpha}(t) = 1. \tag{52}$$

Then for some α and r, $\hat{\Gamma}^r_{\alpha;\alpha} = 1$, $C^r_{i_1,\ldots,i_{n+1}}(t) = 1$, and $A_{i_1,\ldots,i_{n+1},\alpha}(t) = 1$. By induction $A_{i_1,\ldots,i_{n+1}}(t) = 1$, so by induction using (b), $A_{i_1,\ldots,i_n}(t) = 0$, and so $A_{i_1,\ldots,i_n,\beta}(t) = 0$ for all β. But every term of (43) except the first contains the factor $A_{i_1,\ldots,i_n,\beta}(t)$, and hence all these terms are 0. This proves (c).

We do not prove here that the dynamics of A has a limit as $\Delta t \to 0$. To do so requires placing resetting conditions on the rule strength dynamics in such a way that only a finite number of discrete time rules may be selected in succession. A complete analysis of this problem is beyond the scope of this paper.