ELSEVIER

Theoretical Computer Science 217 (1999) 157–172

Theoretical Computer Science

Coupled map lattices dynamics on a variable space for the study of development: A general discussion on *Caenorhabditis elegans*

Franco A. Bignone[a,b,*]

[a] *Dip. Oncologia Sperimentale, Istituto Nazionale per la Ricerca sul Cancro, IST, L.go Rosanna Benzi, 10, I-16132 Genova, Italy*

[b] *Dip. Biologia Animale e Genetica, Univ. di Firenze, Firenze, Italy*

Abstract

The implcmentation of models to simulate dynamical aspects of biological systems, particularly in the study of replication and development, has to take into account that gene–gene interactions and cell–cell interactions take place in a space of variable structure and size. In this paper, we discuss the possibility of a detailed reconstruction of cellular dynamics during early stages of development with coupled map lattices or formalizations of Lindenmayer grammars. © 1999 Published by Elsevier Science B.V. All rights reserved

Keywords: Genetic networks; Gene expression; Caenorhabditis elegans; Coupled map lattices; L-systems; Context-sensitive grammars

1. Introduction

Coupled map lattices (CML), introduced in recent years as discrete models for the study of spatially extended dynamical systems (SEDS), are particularly apt for simulations in the complex situations that arise in Biology. These deterministic models, discrete in time and space, differ from deterministic cellular automata (DCA) in the definition of state variables, these being discrete in DCA and continuous in CML. In this sense CML sit on a middle level between models implementing a traditional formulation through PDEs and DCA.

CMLs have been used to study spatially extended chemical reactions, interface dynamics, biological systems [1, 2, 9, 11]; see [10, 13] and references therein for reviews. Recent numerical simulations using CMLs have shown that their complex dynamics

* Correspondence address. Dip. Oncologia Sperimentale, Istituto Nazionale per la Ricerca sul Cancro, IST L.go Rosanna Benzi, 10, I-16132 Genova, Italy. Fax: +39 10 5600210; e-mail: abignone@igecuniv.cisi.unige.it.

0304-3975/99/$ - see front matter
PII: S0304-3975(98)00155-8

has several characteristics that recall the behaviour of DCA. On a theoretical ground the study of CML has thus given a better understanding on the nature of emergent complex behaviour in SEDS. It has been shown for example that the kind of complex evolution of certain DCA rules, "chaotic" in Wolfram's definition [25], is present in stable CML – stable in the sense of negative Lyapunov exponents. Moreover, the similarity between CML and DCA can be made more clear or through symbolic encoding of CML dynamics, or through the study of CML in which a symbolic representation is built in the model itself. For a discussion about these points see [2–5, 17], for recent experimental evidences on this behaviour in SEDS see [16]. While a description of the biological systems discussed here could be possible using DCA, we decided to implement a model based on CML in order to have a closer connection with experimental situations.

In all the cases reported above the study of SEDS through CML has been limited to dynamics simulated, through parameter changes, on lattices of fixed size and shape from $t(0)$ onwards. The lattice's morphologies have in general been defined over the ring or the torus with periodic or no-flux boundary conditions. In some other cases the dynamics have been analyzed on particular lattice structures, such as the Sierpinski's Gasket, but always with a set of elements constant from the begininng of the simulations [6].

For most biological systems, where cell replication, chemotactic movement, growth, and cell death occurs, this assumption is a limitation in respect of one of the main characteristics shown. Formalization of biological systems as lattices of fixed size can be very important to study their basic dynamical properties. However, for a more realistic study, if we use, for example, a CML formalization to investigate cell–cell interactions, it should be taken into account that local space–time dynamics set the lattice's size and shape, i.e. number of sites and topological distribution of signal sources, while biochemical signals travel through it driving growth, death and/or cell movement.

This phenomenon results in a continuous coupling between the studied dynamics and lattice structure. Biological phenomena that shows this kind of plasticity are known to be quite general and occurs at early and late stages of development, in several pathologies, i.e. cancer progression, and during homeostasis [18, 21].

The different approaches introduced so far to study this problem can be subdivided into two main groups depending on the theoretical tools that have been used for the simulations.

From the point of view of development, with a particular reference to plant morphologies, A. Lindenmayer and co-workers [15, 20, 22] have introduced the definition of L-systems, a particular type of formal grammars with parallel updating, for this purpose. In this respect it should be noted that formal grammars implement in a *natural* way the conceptual structure introduced by biologists in order to describe their experimental systems. In fact, the logical description of development has preceded historically molecular approaches. The history of our concepts about development has founded the description in a way apt for treatment in terms of *computer language* more than in

quantitative mathematical terms. Even the biological concepts of: *fate maps* – i.e. the mapping of prospective parts of the embryo onto the volume of the egg – *developmental hierarchy* – i.e. the hierarchy of decisions taken by cells in order to achieve a certain final differentiative state – *family tree* – i.e. the lineage of derivation of every cell in the body from the initial egg –; all have obvious correspondences with basic concepts in Computer Science.

In recent years approaches more classically oriented from a mathematical point of view, using PDE and geometric considerations, have also been developed in order to study the effect of cell replication in the creation of structure [7].

As a last comment it is important to notice that the theoretical aspects of dynamical systems in which there is a temporal variation in the number of degrees of freedom has been until now poorly investigated, aside from some noteworthy exceptions [11, 12, 15].

In this work we introduce an approach that relies on a general definition, in terms of biochemical genetic networks, of the rules that set a context sensitive, generative, gammar. This model can be seen in turn as a general definition for a class of CML with variable degrees of freedom.

2. C. elegans as an L-system

The biological problem from which we draw our model is the development of Caenorhabditis elegans [C.e.] up to the level of hatching. C.e. is a Nematode worm that has become one of the main experimental systems for the study of development and genetics [26].

A picture of the *four-cell stage* of development is shown in Fig. 1. The egg surroundig the cells has an elliptical shape, and initially contains a single cell, *the zygote*, which by subsequent subdivisions of its mass by cell replication gives rise to 558 fully differentiated cells at hatching in the hermaphrodite; out of a total of 671 cells, 113 of which are eliminated by programmed cell death.

It has been known for a long time that several organisms, from Anellids to Gastropods and Nematodes, have a fixed pattern of division and development. Detailed embryological studies have been described in the past and the pattern of replication in the embryo and genetic studies are in progress in order to define molecular determinants for such behaviour. In Fig. 2 is reported an example of the definitions currently used by embryologists, and the scheme for the first two divisions for the AB lineage are reported. This results in a complex binary tree [26].

The logical scheme followed by embryologists is as follows: besides a few exceptions, initial cells, defined in their characteristics through experimental testing are named by letters, i.e. AB, P2, EMS, the cells originated from these by subsequent cell divisions are named depending on the direction of the division plane. Thus, if a cell divides into two daughters following an anterior–posterior plane, the two daughters cells will be named $AB \rightarrow ABa, ABp$. If the plane is instead on a right–left symmetry line we will use the notation $ABa \rightarrow ABar, ABal$. The cell $ABprpapapp$ results from seven

Fig. 1. Caenorhabditis elegans, early stage of development, after the initial three divisions. Clockwise from the left side, cells are: *ABa*, *ABp*, P_2, *EMS*.

$$
\begin{array}{llll}
S \equiv (ZYGOTE) \rightarrow P_0; & P_0 \rightarrow \{AB\}P_1 & & \\
P_1 \rightarrow \{EMS\}P_2; & \{EMS\} \rightarrow \{MS\}E & & \\
P_2 \rightarrow \{C\}P_3; & P_3 \rightarrow \{D\}P_4 & &
\end{array} \tag{1}
$$

$$
\textbf{Founder cells}: \ \{AB\}, \{MS\}, E, C, D, P_4 \tag{2}
$$

$$
\begin{array}{l}
\{AB\} \rightarrow \{ABa\}\{ABp\} \\
\{ABa\} \rightarrow \{ABal\}\{ABar\} \\
\{ABp\} \rightarrow \{ABpl\}\{ABpr\} \\
\{ABal\} \rightarrow \{ABala\}\{ABalp\} \\
\{ABar\} \rightarrow \{ABara\}\{ABarp\} \\
\{ABpl\} \rightarrow \{ABpla\}\{ABplp\} \\
\{ABpr\} \rightarrow \{ABpra\}\{ABprp\} \\
\ldots\ldots
\end{array} \tag{3}
$$

Fig. 2. Cell lineages in C. elegans represented as a grammar. The distinguished word S correspond to the Zygote, the fertilized egg. Every letter, or groups of letters in curly brackets, represent a cell. The first divisions give rise to the founder cells (1,2), subsequent divisions are in general represented by embryologists by addition of an a or p – in case of an anterior, posterior division – or an l or r for a left, right division, the beginning of the $\{AB\}$ lineage tree is shown (3). Terminal symbols are in this case the final differentiated cells of the replication tree, or cell death '0' in case of apoptosis.

anterior–posterior divisions, of which in five cases has been the posterior daughter, and in two cases the anterior one, and by one left–right division, of which it has been the right cell produced. Thus, the name of a cell contains its history. In the past, other schemes have been used, for example with numbers, but the logic for the description has always been the same or similar.

The development of this organism, from a descriptive point of view, can thus be conceptualized in a very similar way as an L-system as described by A. Lindenmayer [15].

L-systems have been conceived as a formal description of plant development, the emphasis was on plant topology, and by the role played by replication and growth in it. Subsequently, several geometric interpretations of L-systems have been proposed, for example a "turtle geometry" interpretation of L-systems is used in [20]. The relationship between L-systems and formal grammars, such as Chomsky's grammars, lie in the method of applying productions. In Chomsky's grammars productions are applied sequentially, while in L-systems they are applied in parallel.

The general definition of a formal language begins with the introduction of a finite nonvoid set of symbols, a finite alphabet, usually denoted by V. The elements of V are called *letters* or *symbols*, finite strings of letters define words over V. The set of the words over V is denoted by V^*.

A *generative grammar* G is an ordered four tuple (V_N, V_T, S, F), where V_N and V_T are finite alphabets with $V_N \cap V_T = \emptyset$, S is a distinguished symbol of V_N, and F is a finite set of ordered pairs (P, Q) such that P and Q are in $(V_N \cup V_T)^*$ and P contains at least one symbol from V_N. The symbols V_N are called *nonterminal* symbols or *variables* and are denoted by capital letters. The symbols of V_T are called *terminal* symbols and are denoted by small letters. The nonterminal symbol S is called the *initial symbol* and is used to start the derivations of the words of the language. The ordered pairs in F are called *rewriting rules* or *productions* and are written in the form $P \rightarrow Q$.

Keeping this formalism, the development of C.e. can be seen as a generative grammar in which the four tuple (V_N, V_T, S, F) is made respectively of $S \equiv Zygote$, the finite alphabet

$$V_N = \{S, AB, P_0, P_1, P_2, EMS, \ldots ABa, \ldots ABal, \ldots\} \tag{4}$$

will contain symbols that define cellular stages during development, $V_T = \{a_t, p_t, l_t, r_t, 0\}$ or $V_T = \{ABarppappp, ABalaaaalp, \ldots, 0\}$, if one assumes as terminal symbols a special instance of symbols a, p, l, r, as used by embryologists, or instead a description through terminally differentiated cells before hatching. In both cases the term 0 for cell death has to be included, and $F = \{AB \rightarrow ABa, ABp; ABa \rightarrow ABar, ABal; \ldots\}$ the rules of production, applied in parallel as in L-systems. It has to be noticed that, from a biological point of view, the system is not strictly parallel, but for what concern a qualitative description it can be considered as such. The story is obviously different and more complex if one takes this fact into consideration, this possibility will be further discussed below.

The difficult problem to solve is the definition of F, if we do not want just a descriptive approach, but instead out intention is to build a dynamical approach. In fact, for decription it is possible to maintain the scheme above with few improvements, but this is not that much informative, the end result is going to be a clean rationalization of experimental results. Instead if the goal is to be able to define F in terms of local rules of interaction between genes and cells at the molecular level then the problem is quite complex.

3. Definition of a variable CML

Except when studying some simple cases, as for the alga Anabaena catenula [14], the modelling proposed by Lindenmayer has not been connected so far with possible underlying biochemical dynamics driving the system, as is the case for gene-expression or metabolism; see [20].

In order to link, recent results obtained in simulations of spatially extended chemical systems with CMLs, together with L-systems, we propose a model for a parametric context-sensitive L-system defined through gene–cell interactions. The basic model has been proposed recently, through CML formalizations, in order to study a formal scheme of interactions in *trans* among genes, through the distribution in time and space of their products among neighbouring cells [1]. This scheme is a very simplified scheme of gene regulation. Despite its limitations we use here this formalism in order to discuss the overall model. What has to be kept in mind is the fact that modifications in the details of F do not change the general usefulness of this discussion.

Moreover, there is still uncertainty among embryologists if, in the case of C.e., cell–cell interactions do play an important role during development, or if instead cell fate and embryonic development are fully set in an autonomous way. The meaning that embryologist apply to *autonomous* in this case is that of assuming that inner cellular dynamics are solely responsible for the overall behaviour, even if recent experiments seems to show that the real situation is much more complex [23].

3.1. The model

The variation in time of the product of a generic gene G_i in a network of n genes, $i=(1,2,3,\ldots,n)$, as for example a protein x_i inside a certain cell, can be defined as

$$x_i(t+1)=x_i(t)(1-C_i)+v_i(t)P_i, \tag{5}$$

where P_i and C_i represent, respectively, the rate of production and a first-order decay for x_i. The term v_i is the coupling term for the gene network of the single cell, which controls the activity of G_i, influenced by some other genes, and is defined as a discontinuous function as follows:

$$v_i(t)=\begin{cases} 1 & \text{if } \sum_{j=1}^{n} x_j(t) r_{i,j} \geqslant k, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where x_j's are the concentrations of the gene products present in the system, $r_{i,j}$'s are real numbers representing the inter-relationships among them and k represents the threshold value.

Using this simplified description of a cell, the study of the distribution in space of the products, and thus cell–cell interactions, can be done by the addition of a diffusion term in a set of identical equations like the one above, each representing a cell. This procedure gives, for the ith gene of the mth cell of a discrete lattice formed by N identical nearest-neighbour-coupled cells with periodic boundary conditions

$$x_i(m,t+1) = x_i(m,t)(1-C_i) + v_i(m,t)P_i + \gamma_i \left[\sum_{j=1}^{q} x_i(j,t) - q x_i(m,t) \right] \tag{7}$$

the parameter γ above defines the diffusion constant and q the set of the neighbours, e.g. for a ring dynamic with periodic boundary conditions $q = \{m-1, m, m+1\}$.

As was shown previously [1], the isolated map has quite a simple repertoire of behaviour. All steady states, as studied by linear analysis, are stable, and can be single points or oscillations. The length of the period T of the oscillations can be modulated at will through parameters setting and follows a complex bifurcation sequence.

More complex dynamics that are shown by this CML, and by the general class of CMLs in which discontinuous, or very steep, functions are present, have been discussed in several instances [1, 2, 17]. In these models long transients, and long periods, appear with a strict correlation with the strength of the coupling parameter γ, or the structure of the gene network considered. This behaviour is strictly linked to the spatial degrees of freedom present in the system. For a certain genetic network transients and periods shown by these maps are known to grow with lattice size. Moreover, depending on the coupling parameter γ, transients can grow linearly or exponentially [2]. This is an important feature in the case of a growing system. In this case, in fact, once the lattice size reaches a sufficient length – above 40 elements in the worst-case scenario – the final stable attractor, predicted for the isolated map by linear stability analysis, can be computationally unobservable. Thus, the study of long-term dynamics in a system built in this way becomes unobservable for lattice sizes quite small.

The definition of a lattice of variable size, in which cell growth occurs, poses also topological problems. The distribution of the newly formed cells in space imposes a minor rearrangement of neighbouring elements. Moreover, an algorithm has to be set for the birth process.

The most simple approach, if one wants to study dynamical properties, is the dynamics over a ring. The birth of new cells in this case can be taken into account by adding an element in proximity of the mother cell, e.g. on the right, or the left, or through a random choice. Every cell is thus described, in addition to the parameters introduced above, by two pointers setting the neighbourhood. During the simulation, changes in neighbourhood are taken into account by the definition of a q_i specific for every cell, moreover the strict structure of the ring with nearest neighbours is also taken into account by repositioning the pointers for the mother cells.

The determination of the birth process can be based on the characteristic dynamics of the model shown above. Recent studies on the cell cycle have demonstrated the existence of a very complex system of biochemical interactions inside cells giving rise to a series of cyclic events resulting in cell replication. Detailed studies on those biological systems that are presently known in details, such as in the case of *S. cervisiae* or *S. pombae*, have shown that the complex machinery of biochemical interrelations can be reduced to a system of ODE similar to a *reactive media*. In other terms, cells, in their replication cycle, have the possibility to switch from steady states that are: fixed points, limit cycle oscillations, or an excitable cycle [24, 8]. In case the cells are in active proliferation the oscillatory limit cycle is the case in point. In this respect the model described here, which shows oscillatory steady states, can be considered a simplified, crude, approximation of such behaviour. We are thus limiting ourselves to the case in which cells are actively proliferating.

The rationale behind this crude assumption is moreover due to two additional reasons: (a) the biochemical interactions in higher organisms are much more complicated than those shown by *S. cervisiae* or *S. pombae*, this implies that a jump to the level of a fully realistic model in biological terms is out of question, until some fundamental theoretical work has been done in order to clarify the basic dynamics, (b) the network of biochemical interconnections which regulate cell–cell interactions is very complex and to a large extent still unknown, it is thus better to study a general system with a wide range of parameter values in order to get a blurred picture of the general possible behaviour of the dynamics, and fill in details later, once this knowlegde is acquired.

If one fulfills these assumptions a particular set of concentrations of the variables at one point of the cycle, period, become the switching system for the birth process. In this way, if one consider the coupled model, it is possible to study in a simplified way how changes in the number of degrees of freedom can influence the dynamics in respect of a lattice of fixed size.

For what concerns the transient, typical behaviour for the case of a period $T=3$ attractor for the isolated network and for a long transient over the lattice dynamic is shown in Fig. 3. In the figure the local concentrations for the variables x_i, in a model with two genes, for two sites of the network are plotted. As can be seen there is a dispersion of the values caused by the disturbances travelling through the coupled system which ultimately give the long term dynamics observed.

One can thus establish, similar to what is thought to happen in vivo, that a certain range of concentrations, is able to switch replication. A further set of parameters: $l_i(min), l_i(max)$, set the range limits that trigger such behaviour, for any number of gene products in the network. The biological assumption behind this idea is that these concentrations will trigger some cascade of events leading to replication.

Results obtained with the dynamics of rings starting from random initial conditions for a ring of $N=20$ elements, and allowing the ring to grow up to $N=1000$ elements, or for $t=1000$ time steps, are shown in Figs. 4 and 5. In all these cases the model was started from random initial conditions for x_i. The parameter value for the replica

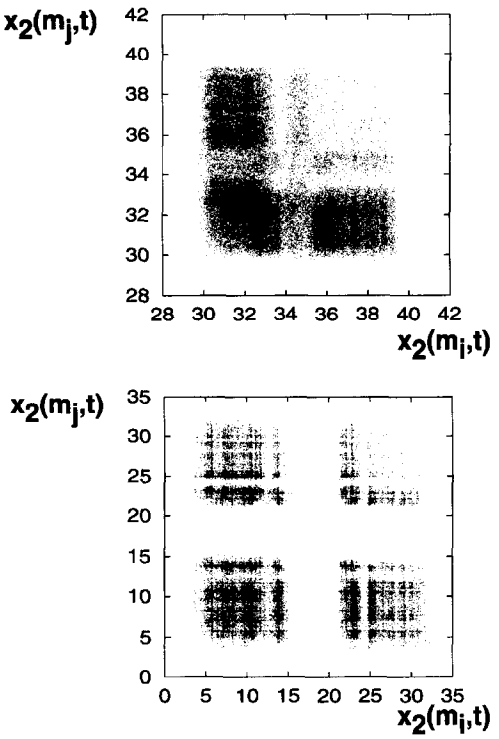

Fig. 3. Simulation of a CML dynamics over the ring. Dispersion of the values of the concentrations of the gene, x_2, in a network of two genes in two sites of the lattice, m_i, m_j, over the lattice during a dynamics over the ring. Above: $\mathbf{R} = [-1, 1, 1, -1]$, $C_1 = 0.2$, $C_2 = 0.4$, $P_1 = 3.5$, $P_2 = 19.0$, $\gamma_1 = 0.2$, $\gamma_2 = 0.7$. Below: $\mathbf{R} = [0, 1, 1, -1]$, $C_1 = 0.1$, $C_2 = 0.05$, $P_1 = 3.3$, $P_2 = 5.0$, $\gamma_1 = 0.0$, $\gamma_2 = 0.9$. Lattice size $N = 200$, $T = 3$ for the isolated map in both cases.

of the network are set at the beginning of the simulation, and left unchanged during the run.

Depending on the choice of the swithching threshold, from the parameters for the isolated map, and from the coupling, different dynamics arise, with different speeds of growth, see Figs. 4 and 5 for example. In the cases shown the system is unbounded, i.e. there are no limits set to the growth. This implies that the long-term dynamics can be studied only for parameter values that increase the number of elements very slowly, in all the other cases the long-term dynamics is unobservable.

In the case of two- or three-dimensional dynamics the topological problem is obviously more difficult to solve. Lattice growth in this case has to take into account the presence of a more complicated set of interactions surrounding every cell.

The possible solutions in this case are quite different depending on the use of the model. In the case of studies related to the analysis of the dynamics of the system a mixed approach similar to that implemented with lattice gas cellular automata (LGCA) with CMLs that move reciprocally can be used. LGCA are models introduced in order

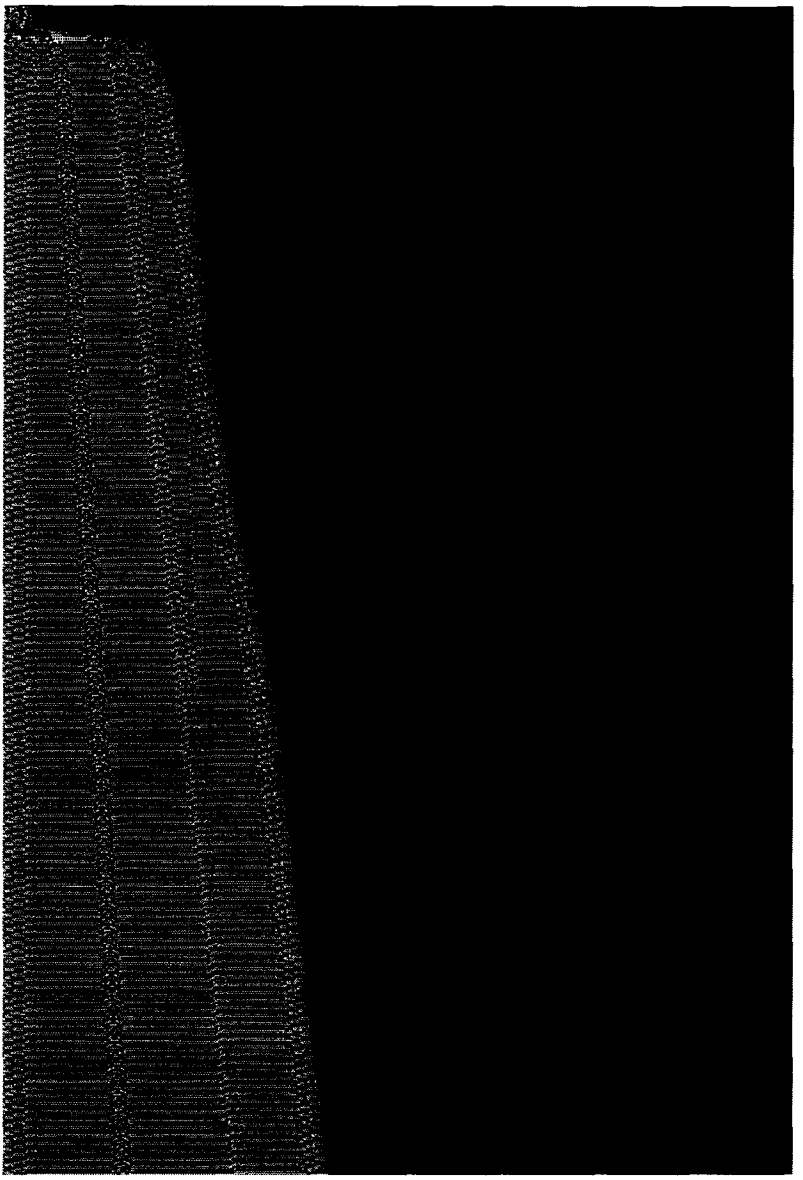

Fig. 4. The plot shows a growth dynamics over the ring, the periodic boundary conditions. Every cell is represented as a picture point, located on the abscissa following m, and colour coded depending on the concentration of one of the genes x_i – only one of the genes is shown –. The time course of the simulation goes from top to bottom. See text for further explanations.

to study fluid flow turbulence, and that have been applied recently to the study of chemical reactions; see [13] for further references and applications to reactive LGCA. In this case the models simulate particle dynamics assuming a discretization for space, time and particle's velocity.

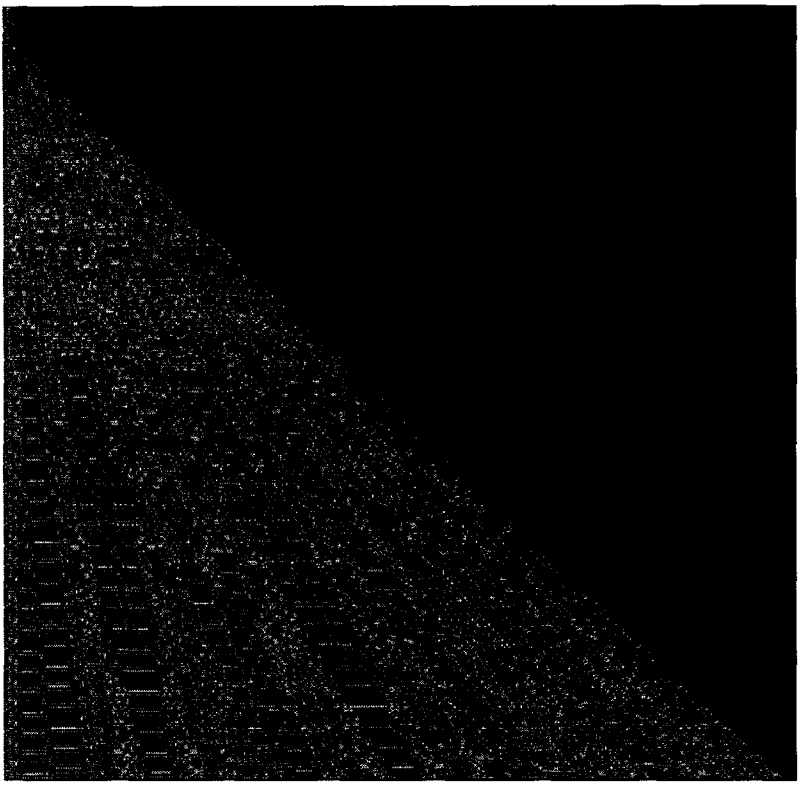

Fig. 5. Growth dynamics over the ring, the plot has been done as in the previous picture, but with a different set of choices for parameters. See the previous caption and text for further explanations.

In order to simulate cellular movement one can build similar models where maps move in space following rules similar to those outlined in LGCA. In this case cells sit on a grid where they can move, and on which they can replicate. Further possibilities of simulation can be implemented using CAs or PDEs, as was reported in the introduction.

4. CMLs as L-systems

We now introduce the last step in modelling this system, the connection between the behaviour of the CML and the description in embryological terms through an L-system. In this context we have to refer to parametric OL-systems. These operate on *parametric words*. In this case the alphabet V is made of *modules* which include a letter, e.g. $X \in V$, and parameters $x_1, x_2, x_3, \ldots, x_n \in \mathbb{R}$ associated with X.

A parametric OL-system is thus defined as an ordered quadruplet $G = \langle V, \Omega, S, F \rangle$, where V is the *alphabet* of the system, Ω is the set of formal parameters, S is a nonempty word and F is a finite set of productions.

A further complication necessary to keep with the biological description implies the introduction of local interactions, the system thus becomes a parametric OL-system that is context-sensitive and for such a case the productions have to be of the type

$$X_{m-1}(x_1,\ldots,t)X_m(x_1,\ldots,t)X_{m+1}(x_1,\ldots,t) \rightarrow X_{m-1}(x_1,\ldots,t+1)X_m(x_1,\ldots,t+1)X_{m+1}(x_1,\ldots,t+1) \quad (8)$$

in which the state of the $X_m(\ldots t+1)$ element is determined by its parameters at time (t) and by its neighbours.

We can now rephrase our CMLs in the following way: we define a generative grammar context sensitive, with parallel updating, in which the alphabet V is defined by $V=\{00,11,01,10\ldots\}$, i.e. we define the alphabet as the possible states of our genetic networks corresponding to $v_i(t)$ in Eq. (6), the numerosity of our alphabet will be defined by 2^n where n corresponds to the number of genes present in the isolated cell, S corresponds to an initial set of genes, i.e. $S=\{001,111,001,110,001\}$ for a group of five cells with three genes each, while $\Omega=\{x_i,r_{ij},v_i,k_i,l_i(max),l_i(min),C_i,P_i,\gamma_i\}$ is the set of parameters, as defined in Eq. (7), and F the productions. The productions will be set following Eq. (7) but including rules for replication of the type

$$X_{m-1}(001,x_i\ldots)X_m(110,x_i\ldots)X_{m+1}(110,x_i\ldots) \rightarrow X_{m-1}(001,x_i\ldots)X_m(110,x_i\ldots)X_{daughter}(110,x_i\ldots)X_{m+2}(110,x_i\ldots). \quad (9)$$

We can define in this way functional states of cells, or cells in cell lineages in terms of gene activation. The organism is thus defined in its components by the differentiative states, or embryonal states, of its cells. Once a set of arbitrary defining words V^* is established the resulting language L, which includes cell states, will define the organism, and some of its dynamics, in a formal manner.

5. Early stages of development and cleavage

The description of development or cell growth using the model described above is obviously quite simplified. Moreover, it is mapped correctly on space only in the case in which cell replication gives rise to net increase of dry mass, i.e. increase in number of cells with an identical volume.

In a lot of animals, cell divisions at early stages are typically cleavage divisions. The number of cells in the embryo will always increase, but cell divisions will always produce daughter cells with a size that is half that of the mother. The case is different for viviparous animals, or those animals in which the egg subdivides early in two regions: embryonic and extra-embryonic. This latter part contains in general a lot of yolk, and the embryo will grow taking advantage of the yolk mass.

Thus for those animals whose eggs do not contain a large quantity of yolk, the model, as discussed so far, does not map space correctly. For example, in the case of C.e., the final number of cells is obtained by subsequent subdivision of the initial

zygote with a reduction in size of about 5×10^2. In this case there is no increase in dry mass, and the animal will grow only after the egg is hatched.

The modelling proposed in the case of this animal can still be interpreted as being symbolic, but in this case there is no connection between our CML's representation and *real* space. Another possibility is to increase the complications of the system. In this case we can keep the correct representation of the CML for the initial space. We will thus have a certain number N of initial sites that at t_0 will all belong to S. While the process of division goes on, the constant number N of sites will keep on being subdivided among the different cells-words defined by the dynamics.

The new definition would be the same for V and S as above, for what concern Ω we will have to introduce further variables

$$\Omega = \{x_i, r_{ij}, v_i, k_i, l_i(max), l_i(min), C_i, P_i, \gamma_i, \beta_i, q_a, q_b\} \tag{10}$$

where γ_i represents diffusion among sites of the CML that belong to the same cell and β_i diffusion between sites that belong to different cells.

In this case the productions would still be of the type

$$\begin{aligned} &\mathbf{X}_{n-1}(001, \mathbf{x}_i \ldots)\mathbf{X}_n(110, \mathbf{x}_i \ldots)\mathbf{X}_{n+1}(110, \mathbf{x}_i \ldots) \\ &\quad \rightarrow \mathbf{X}_{n-1}(001, \mathbf{x}_i \ldots)\mathbf{X}_n(110, \mathbf{x}_i \ldots) \\ &\qquad \mathbf{X}_{\text{daughter}}(110, \mathbf{x}_i \ldots)\mathbf{X}_{n+2}(110, \mathbf{x}_i \ldots) \end{aligned} \tag{11}$$

with a vector, or matrix representation that keeps track of space subdivision. The elements of the CML would be thus of fixed size, but mapped on a description of a grammar where the size and number of elements is variable. In this case the label n does not correspond to the label m mapping the CML as above. For the calculations we will have to keep into account the different neighbours q_a and q_b for every element of $\mathbf{X}_n$; moreover, the application of the regulation must be done only on one site for every $\mathbf{X}_n$, assuming a single location for the regulated gene. In the same way production and catabolism can be applied in a structured manner.

As a final complication an algorithm must be set in order to subdivide space after a replication occurs, assigning to every element $\mathbf{X}_n$ a set of the maps available on the lattice sites that belonged to the mother. This model results in a more accurate description of the internal cellular dynamics with respect to the simplified version, and more correct for the representation of development when cleavage occurs.

6. Results and discussion

The study of dynamical models with a variable number of elements is still an open field. This implies that quite a large amount of preliminary work is necessary in order to study these problems from the perspective of dynamical systems theory. In this respect the model, grounded on an experimental basis derived from the study of

biological systems, is simple enough to allow a detailed theoretical study of its long-term dynamical behaviour.

From the point of view of the study of biological systems, the model presented here is particularly suited for the study of development, adopting a very general point of view. Because of this fact it can be used as a tool to validate working hypothesis in a broad range of different situations, or as a conceptual frame for rationalization in the case of data collections.

In the case of C.e. its usefulness is based on the fact that: cell lineages, i.e. the detailed pattern of cell replications, cell–cell interaction, expression of specific genes that set embryonic development, and cellular positions, through time-lapse cinematography studies, are all available in this particular model. This means that it is possible to trace all the lineages with cell divisions, cell number, and cell movement. This gives rise to a complex structure of a dynamic binary tree, in three dimensions, representing development. Theoretical modelling of this behaviour and testing known hypothesis based on genetic findings, is now possible because of this particularly useful setting. Moreover, results obtained through simulations, owing to the possibility of genetic manipulations, can be tested through experiments also in their prediction capability. The full picture of C.e. development is still being worked out, but has now reached a critical mass already amenable for modelling.

At the same time, because of the huge complications present in biological systems, a word of caution is obviously necessary. The implementation of particulars, in terms of molecular interactions, will obviously require sets of different models, in order to understand fully the dynamics of the possible interactions inside the system. This will pose most likely new questions, one of what is to what extent will we be able to pursue such analysis.

For example, all variables described in

$$\Omega = \{x_i, r_{ij}, v_i, k_i, l_i(max), l_i(min), C_i, P_i, \gamma_i\} \tag{12}$$

are complex variables in vivo; the same holds for topological variables, not discussed extensively here. This implies that there is quite a large amount of work for a good experimental and theoretical definition. At the same time, an ensemble view is important. The main reason being the possibility to get a set of general considerations drawn from the type of modelling described here. The approach that one has to follow to obtain this result is to study separately the problems present in the different subsets knowing complications that can arise out of the global picture.

As was already mentioned this work has also quite a general value, the extension to other systems is obviously possible following a similar rationale. A big problem in perspective is the fact that other experimental models at the moment do not allow such a detailed description because of the intimate nature of their dynamics, or because of their size.

At the same time the study of these simplified situations can give good hints both for further theoretical studies and for experimental work. In order to clarify this point an example can be given, based on simulations done so far. If one considers dynamics

over the ring for the CML described in (7) the isolated map is stable by construction, this mean that the only possible states for the isolated maps can be stable points or oscillations. Despite this 'simple' behaviour the coupling among elements of the CML can give rise to long-term dynamics with transients that increase with lattice size, the system is thus bound to settle into two different main modes of behaviour in terms of long term dynamics depending on the coupling γ_i [2].

For a certain subset of coupling strengths the final steady state is reached in a relatively short time, while for a certain set of values of γ_i transients grow exponentially with lattice size. For systems of a sufficiently large size there is thus the possibility of locking into a never-ending transient dynamics before the final steady state is reached. This is obviously suggesting an interesting possibility of regulation for biological systems where the local coupling of large cell populations is the norm [2].

Another interesting aspect in the definition of γ_i, and in these aspects of transient dynamics, is the fact that in vivo the transmission of information, inside cells and among them, is done through a complex set of biochemical interactions generally indicated as "*signal transduction pathways*". The high complex molecular interrelations present in these systems could drive the dynamics in different regions of parameter space modulating in this way the transients. Local signal transmission can thus by itself give rise to quite a complex pattern of behaviours, aside from other possible mechanisms. The result of this modelling is thus the build up of a rationale where, the different aspects of cell–cell interactions that are initially studied separately, can be linked together in order to understand their complex dynamics.

Acknowledgements

The author wishes to thank Roberto Livi, Antonio Politi, Salvatore Di Gregorio, Ralf and Heinke Schnabel for helpful discussions. Thanks to Ralf Schnabel for Fig. 1. Simulations were done on a Personal IRIS, Silicon Graphycs Inc., HP740, Hewlett Packard Inc., and IBM-RISC6000, IBM Inc. Figs. 4 and 5 have been obtained through data formatting using the *ximage HDF* (Hierarchical Data Format) package, from the National Center for Supercomputing Applications at the University of Illinois in Urbana Champaign, Illinois, U.S.A. This work has been partially supported by a C.N.R. grant #95.01751.CT14 "Studio analitico della dinamica della regolazione genica e della morfogenesi".

References

[1] F.A. Bignone, Cells-gene interactions simulation on a coupled map lattice, J. Theoret. Biol. 161(2) (1993) 231–249.

[2] F.A. Bignone, R. Livi, M. Propato, Dynamical stability and finite amplitude perturbations in coupled genetic-networks, Physica D 108 (1997) 379–396.

[3] F.A. Bignone, R. Livi, M. Propato, Symbolic and global dynamics in coupled genetic networks: evolution of artificial cell populations at the edge of and within chaotic transient behaviour, Proc. 4th Eur. Conf. on Artificial Life, ECAL '97, MIT Press, Boston, 1997.
[4] F.A. Bignone, R. Livi, M. Porpato, Complex evolution in genetic networks, Europhys. Lett. 40 (1997) 497–502.
[5] J.P. Crutchfield, K. Kaneko, Are attractors relevant to turbolence?, Phys. Rev. Lett. 60 (1988) 2715–2718.
[6] M.G. Cosenza, R. Kapral, Coupled maps on fractal lattices, Phys. Rev. A 46 (4) (1992) 1850–1858.
[7] S. Duvdevani-Bar, L. Segel, On topological simulations in developmental biology, J. Theoret. Biol. 166 (1994) 33–50.
[8] A. Goldbeter, A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase, Proc. Natl. Acad. Sci. U.S.A. 88 (1991) 9107–9111.
[9] A. Hastings, K. Higgins, Persistence of transients in spatially structured Ecological Models, Science 263 (1994) 1133–1136.
[10] K. Kaneko (Ed.), Overview of coupled map lattices, CHAOS 2, 1992.
[11] K. Kaneko, T. Yomo, Cell division, differentiation and dynamic clustering, Physica D 75 (1994) 89–102.
[12] K. Kaneko, T. Yomo, Isologous diversification: a theory of cell differentiation, Bull. Math. Biol. 59 (1) (1997) 139–196.
[13] R. Kapral, Discrete models for chemically reacting systems, J. Math. Chem. 6 (1991) 113–163.
[14] C.G. Koster, A. Lindenmayer, Discrete and continuous models for heterocyst differentiation in growing filaments of blue-green bacteria, Acta Biotheoret. 36 (1987) 249–273.
[15] A. Lindemayer, Mathematical models for cellular interaction in development, Parts I and II, J. Theoret. Biol. 18 (1968) 280–315.
[16] V. Petrov, Q.I. Ouyang, H.L. Swinney, Resonant pattern formation in a chemical system, Nature 388 (1997) 655–657.
[17] A. Politi, R. Livi, G.L. Oppo, R. Kapral, Unpredictable behaviour in stable systems, Europhys. Lett. 22 (1993) 571–576.
[18] R.T. Prehn, Many growth factors may not be growth factors, Cancer Res. 52 (1992) 501–507.
[19] P. Prusinkiewicz, J. Hanan, Visualization of botanical structures and processes using parametric L-systems, in: D. Thalmann (Ed.), Scientific Visualization and Graphic Simulation, Wiley, New York, 1990, pp. 183–201.
[20] P. Prusinkiewicz, A. Lindenmayer, The Algorithmic Beauty of Plants, Springer, Berlin, 1990.
[21] M.C. Raff, B.A. Barres, J.F. Burne, H.S. Coles, Y. Ishizaki, M.D. Jacobson, Programmed cell death and the control of Cell survival: lessons from the nervous system, Science 262 (1993) 695–700.
[22] G. Rozenberg, A. Salomaa, The Mathematical Theory of L Systems, Academic Press, New York, 1980.
[23] R. Schnabel, Autonomy and nonautonomy in cell fate specifications of muscle in the Caenorhabditis elegans embryo: a reciprocal induction, Science 263 (1994) 1449–1452.
[24] J.J. Tyson, Modeling the cell division cycle: cdc2 and cyclin interactions, Proc. Natl. Acad. Sci. U.S.A. 88 (1991) 7328–7332.
[25] S. Wolfram, Cellular automata as models of complexity, Nature 311 (1984) 419–424.
[26] W.B. Wood, The Nematode Caenorhabditis Elegans, Cold Spring Harbour Laborataries, New York, 1988.