INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

Transition to Chaos in Models of Genetic Networks

Colin C. Hill

Centre for the Physics of Materials, Department of Physics
and
Centre for Nonlinear Dynamics, Department of Physiology

McGill University
Montréal, Québec
Canada

A Thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Master of Science

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

0-612-44182-2

This thesis is dedicated to the memory of
Carl Sagan whose book "Shadows of Forgotten Ancestors",
co-authored with his wife Ann Druyan, aroused my interest
in the biological world.

"When we no longer look at an organic being as a savage looks at a ship, as something wholly beyond his comprehension; when we regard every production of nature as one which has had a long history; when we contemplate every complex structure and instinct as the summing up of many contrivances, each useful to the possessor, in the same way as any great mechanical invention is the summing up of the labour, the experience, the reason, and even the blunders of numerous workmen; when we thus view each organic being, how far more interesting–I speak from experience–does the study of natural history become!"

CHARLES DARWIN,
The Origin of Species
(appearing in "Shadows of Forgotten Ancestors")

Contents

Abstract

Randomly constructed networks of N elements governed by piecewise linear differential equations have been proposed as models for neural and genetic networks. In this model an element is labelled "on" if it is above a threshold, and "off" otherwise. The logical state of the network at any time is thus a Boolean vector. For each element, there is a rule (truth table) specified by the values of K input elements that determines whether it will switch its state (from 1 to 0 or from 0 to 1) at some future time. Parameter p biases the output state of each automata and influences the global dynamics of the system. Discrete switching network analogues of the piecewise linear differential equations show a transition from ordered to disordered dynamics as p and K are varied. The transition is analyzed in continuous switching networks. Previous studies of these networks have demonstrated the existence of steady state, periodic, and chaotic attractors. Numerical studies analyze the dynamics of a randomly chosen ensemble of networks as a function of N, K, and p. Theoretical insight into the change in dynamics is gained through the derivation of a Markov process for the number of elements at time t that are approaching their threshold. Analysis of this Markov model yields a critical relation between p and K that separates ordered dynamics from deterministic chaos.

Résumé

Des réseaux aléatoires de N éléments régis par des équations différentielles linéaires par morceaux ont été proposés comme modèles de réseaux génétiques et neurologiques. Dans ce modèle, un élément est actif (ON) s'il dépasse un seuil, inactif (OFF) s'il lui est inférieur. L'état logique du réseau est ainsi, en tout temps, un vecteur booléen. Pour chaque élément, une table de vérité spécifiée par les valeurs de K éléments d'entrée détermine si cet état est modifié (passant de 0 à 1 ou de 1 à 0) dans le futur. Un paramètre p biaise la sortie de chaque automate et influence la dynamique globale du système. Les analogues discrets des équations différentielles linéaires par morceaux montrent une transition d'un état ordonné à un état désordonné lorsque les paramètres p et K varient. Cette transition est étudiée dans les résaux continus. Des études antérieures de ces réseaux avaient mis en évidence l'existence d'états stationnaires, d'attracteurs périodiques de même que d'attracteurs chaotiques. Des simulations numériques permettent de comprendre l'influence des paramètres N, K et p sur un ensemble aléatoire de réseaux. Ces changements dynamiques sont interprétés a l'aide d'un processus de Markov décrivant le nombre d'éléments approchant leur seuil. Une analyse de ce modèle markovien montre un lien critique entre les valeurs des paramètres p et K séparant les zones de dynamique ordonnée des régions chaotiques.

Acknowledgments

My involvement with research on genetic networks began out of my interest in the question of how life originated. According to accepted scientific dogma, the universe began in a big bang about 20 billion years ago. Elements formed, then eventually planets. In 1953 Stanley Miller showed that it is possible to spontaneously synthesize the complex amino acids which form the basic building blocks of life from the simple elements of carbon, nitrogen, oxygen, and water vapour. Darwin's theory of natural selection attempts to account for how all higher organisms evolved from simpler organisms. So according to the ideas of evolution, all life on earth evolved from one or a small number of original living entities. The gap in the story of how we came to be exists between the formation of amino acids and the emergence of the first living entity. This is the problem that brought me to the Santa Fe Institute to work with Stuart Kauffman on his model of the origin of life. However, once I got to New Mexico I was persuaded to get involved in research on random Boolean networks that model the regulatory genetic network giving rise to cell differentiation.

I would like to thank everyone who contributed to this work. I have sought out the help and support of many individuals in an effort to complete the work embodied in this thesis. I would most like to thank my supervisor Leon Glass for his help and guidance throughout this project. His flow of ideas and insights kept my wandering gaze focused on the key issues related to our work. I would also like to thank him for formulating the differential equation analogues of the discrete logical switching network that is the basis of this work. I would like to thank Martin Grant who was my co-supervisor during most of the year I spent at McGill. His willingness to discuss problems related to this project was very helpful. I would like to thank Stuart Kauffman for getting me started on this research. I also thank him for monetary support during the summer of 1996 and for inventing the model of the genetic regulatory system that bears his name. I would like to thank Royce Zia and Beate Schmittman for getting me started on the research of complex non-equilibrium systems. I would like to thank my family; Foster, Joan, Virgil, and Bonilyn Hill for their love and support throughout the last couple of years. I would especially like to thank my mom for being devoted to me and my efforts. Hearing her voice on the telephone provided more encouragement to me than she will ever know.

I would also like to thank collaborators Bruce Sawhill and Thomas Mestl for many helpful discussions and for initiating some of this work. I want to thank two of my closest friends, Mikko Haataja and Jeremy Taylor for letting me talk about the 'deep, but uninteresting' questions of science through many a lunch hour. I also thank them for helpful discussions and ideas directly related to this research and for reviewing old versions of this manuscript. I thank Kevin Hall and Yoshihiko Nagai (Tiger) for their assistance in the preparation of the manuscript and for helpful discussions. I thank Bhavanna Khanna and Joel Ariathnam for their interest in my work. I thank Josh Pomeroy, Mikko Karttunen, Peter Alleyne, Gianni Taraschi, Tim Carlson, Lolly Brown, and Nick Cabatoff for assistance in computing and software matters. I would also like to thank Caroline Rousseau for translating my English abstract to French. I thank the Santa Fe Institute for accepting me to the 1996 Complex Systems summer school and for use (overuse) of their computing resources. I am indebted to NSERC for their support of this research and my education during my stay in Montreal. I thank the McGill Dept. of Computer Science for use of their computers. I also thank the McGill Dept. of Physics, the Center for Nonlinear Dynamics in Physiology and Medicine, and the Cornell Physics Department for providing resources and a home to engage in research. I also thank Lisa Godwin, Dude, Sajan Saini, Claude Theoret, Diane, Tiago Dejesus, Jamie Cohen, Richard Hill, Cotton Randall, Oleh Baran, Danny Helmick, Joe Cahill, Nellie Tungteng, Veirke Tungteng, Suiru Tungteng, Anne-Marie Lanctot, Claire Riley, Chris Moore, Bill Macready, Jenny Zia, Debra Hatfield, Doug Fitchen, Eberhard Bodenschatz, Jim Sethna, Zoltan Torockzai, Andy Wuensche, Andy Goldenson, Jonathan Taylor, Mary Elle Carpenter, Brenda Barclay, Judith Mueller, Leah Bonema, Erol Girt, Mohsen Sabouri, and my Danish office mates, Louise Kold Andersen, Trine Krogh Madsen, and Dorte Skriver.

Abbreviations and Symbols

$K \equiv$ number of inputs to each element
$N \equiv$ number of elements in the network
$p \equiv$ probability of having a 1(0) in the output of a rule table
$X_i \equiv$ Boolean representation of element i
$x_i \equiv$ continuous, real number representation of element i
$\vec{X} \equiv$ N-dimensional vector of the Boolean configuration of the system
$\Lambda_i \equiv$ randomly chosen Boolean function
$\vec{X}_{(i)} \equiv$ K-dimensional vector of the inputs to an element
$t_{tr} \equiv$ length of the transient before an attractor is reached
$T \equiv$ period of a limit cycle

1

Introduction and Review

1.1 Introduction

Logical switching networks are dynamical systems that evolve in discrete time steps. They are composed of elements that assume discrete values and each element changes its state according to a logical function of the values of its inputs. The idea that complex biological control systems could be modeled as logical switching networks was first developed by McCulloch and Pitts in 1943 [1]. McCulloch and Pitts introduced the concept of the 'formal neuron' as a binary threshold element. They demonstrated that a collection of these elements, 'a neural network', could act as a universal Turing machine [2]. In 1969, Kauffman introduced a logical switching model of the regulatory genetic circuits which gives rise to cellular differentiation during embryonic development [3]. The state of activity of a gene is 0 or 1, corresponding to a gene being repressed or expressed. The state of a gene at a given time step is a Boolean function of the values of K input genes at the previous time step. Elements are updated synchronously. There are 2^K possible combinations of K Boolean elements and thus 2^{2^K} possible Boolean functions for updating the state of a given gene. The unknown nature of the complex interactions taking place in a genome led Kauffman to assume that these could be modeled by randomly chosen functions of randomly selected genes. These choices are then kept fixed for the evolution of the system. For N elements, there are 2^N possible states which implies that the trajectory of a system through state space eventually becomes periodic within a maximum of 2^N time steps. Relevant observables in such a randomly constructed dynamical system are the are the number of cycles and their length. Kauffman proposed this model for cellular differentiation by interpreting the different cycles as different cell types and cycle lengths as cell cycle times [3, 4, 5]. Early numerical results indicated that for $K = 2$, the length and number of cycles $\sim \sqrt{N}$. This relationship did not hold for other values of K. Kauffman argued that $K = 2$ networks are biologically significant because the number of cell types and cell cycle times of eukaryotic organisms is proportional to the square root of the number of genes in their DNA, over a wide range of phyla [3]. The Kauffman model has attracted a great amount of attention from theoretical physicists [6-19]. Concepts from statistical mechanics useful in the study of phase transitions and critical phenomena were employed to understand the qualitative change in global dynamics as parameters are varied. The first effort was made in the context of a modified version of the model which was more accessible to analytic study than the original model. This was the 'annealed' Kauffman model, where the inputs and Boolean rules are no longer 'hard-wired' or 'quenched', but are re-selected at each time step. Analytical work related the convergence or divergence of different initial configurations to the change in global dynamics [7]. The second major effort identified the percentage of frozen elements as an order parameter. Changes in this quantity were correlated with the change in the global dynamics [11].

Binary variables and synchronous updating used in discrete random Boolean networks (RBNs) are an idealization of more biologically realistic continuous differential equations where the state of activity of a given gene is a continuous variable and time is continuous. In 1975, Glass introduced coupled piecewise linear differential equations as a continuous version of the Kauffman model [20]. The logical structure and connectivity in the continuous model is constructed in an identical manner as in the discrete systems. However, in this model, the state of a gene is represented by a real number, x_i, with a corresponding coarse grained Boolean variable X_i that is defined as follows:

$$X_i = 0 \text{ when } x_i < 0;$$

$$X_i = 1 \text{ when } x_i > 0.$$

$$X_i = 1 \text{ when } x_i > 0.$$

In general, only one element crosses its threshold at a time [20]. The Boolean configuration of the system (i.e. the state of the N discrete Boolean variables, $\{X_i(t); i = 1, 2, ..., N\}$) can therefore only change by one element at a time so that the evolution of the discrete configuration can be mapped onto flows along the edges of an N-dimensional hypercube [20]. The asymptotic dynamics are steady states, limit cycles, quasiperiodicity, and deterministic chaos [21, 22, 23].

There is a parameter in addition to N and K that affect the dynamics of both the discrete and continuous systems: the bias towards 1 or 0 in the output of the rule table of each element, designated as p. Recent work on continuous networks has demonstrated chaotic dynamics [21, 22, 23, 24], but the transition from steady state and limit cycle dynamics to chaos has not been studied.

This present work analyzes the dependence of the proportion of chaotic attractors in an ensemble of continuous networks on p, K, and N. The thesis is organized as follows. In the introductory chapter, results from studies of discrete and continuous RBNs which have direct relevance to this current study are reviewed. Chapter 2 is based on the manuscript "Ordered and Disordered Dynamics in Random Networks" which describes the transition from steady state and limit cycle dynamics to deterministic chaos as p is tuned. This manuscript was written with my supervisor Leon Glass and was published in Europhysics Letters, **41** (6), pp. 599-604 (1998). It contains the results of numerical simulations and theoretical predictions of dynamics in continuous RBNs as a function of p, K, and N. These results are summarized and discussed in Chapter 3, along with future directions for this work.

1.2 Discrete Switching Networks

1.2.1 Definition of Model

In the Kauffman model, the genome or network is represented as N genes or elements that can have one of two values at a given time t: $X_i(t) = 0, 1$; $i = 1, 2, ..., N$. A gene is dormant if $X_i(t) = 0$ and is active if $X_i(t) = 1$. The state of the i^{th} element of the network at time step $t+1$ is determined by a randomly chosen Boolean function, $\Lambda_i(\vec{X}_{(i)}(t))$, of the state of K randomly chosen input elements at time t, $\vec{X}_{(i)}(t) = \{X_{j_1(i)}(t), X_{j_2(i)}(t), ..., X_{j_K(i)}(t)\}$. There are 2^K possible combinations of the values of the K inputs and 2^{2^K} possible Boolean functions. The number of possible rule combinations in a network is therefore $2^{N\times 2^K}$. There are $\begin{pmatrix} N \\ K \end{pmatrix}$ ways of choosing the K inputs for each element, and thus $N^{\begin{pmatrix} N \\ K \end{pmatrix}}$ ways of constructing the connection topology of the network. The number of possible realizations of a network $\sim N^{\begin{pmatrix} N \\ K \end{pmatrix}} 2^{N\times 2^K}$. Even for modest values of N and K, this number is astronomically large. Previous work has analyzed dynamical properties of networks (i.e. number of cycles, median cycle length, etc.) for fixed values of N and K by sampling random networks with fixed N and K [3]. These networks can generate widely varying dynamical properties which necessitates studying a large sampling of networks.

The time evolution of each site is given by

$$X_i(t+1) = \Lambda_i(\vec{X_{(i)}}(t)); i = 1, ..., N. \tag{1.1}$$

As the variables in this network take on only discrete values, the number of configurations in state space is finite and implies that the dynamical evolution of this system must eventually become periodic. An asymptotic periodic orbit is designated as a cycle and the transient time until the trajectory reaches a cycle is designated as t_{tr}. Since the number of states available for the system to explore increases exponentially with N, it is possible to have very long transients and cycles. There can also exist a

A	B	C	$i(t+1)$
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

Table 1.1: Typical rule table of an element for K=3 and p=0.5.

large number of different attractors. Initial configurations that lead to a particular attractor are said to belong to the *basin of attraction* of that particular attractor.

In this model, the probability of choosing a 1 in the output of a rule table for each element is p. In the original Kauffman model, $p = 0.5$. A typical rule table for $K = 3$ and $p = 0.5$ is shown in Table 1.1. The influence that p has on the dynamics of the system can be easily seen in the extreme case of $p = 1$, where the rule table of each element would contain all 1's. In this case, each element of the network would be fixed to 1 (i.e. a steady state attractor) after one time step. In general, increasing p affects the dynamics by giving elements an increased tendency to be fixed to 1.

The relevant questions to ask about the properties of such a randomly constructed dynamical system are, for given values of N, K, and p:

- What is the average or median number of attractors?
- What are the expected numbers of steady states and cycles?
- What is the expected length of the transients and the periods of cycles?
- How many initial configurations belong to the basin of attraction of a particular attractor?

- Are the properties of attractors and transients dependent on whether one is averaging over different initial configurations of a given network or averaging over different networks?

- How are the dynamics of a network related to the structure of the network?

1.2.2 Results

The earliest work of Kauffman [3, 4, 5] attacked some of these questions numerically. He found that the number and median cycle length of the attractors were proportional to $\sqrt{N}$ for $K = 2$, $p = 0.5$ networks. Log-log plots of the median cycle length and median number of cycles are shown in Figure 1.1 [3]. For networks with $K = N$, the expected cycle length is $2^{N/2}$. Kauffman argued that the short cycle lengths and small number of attractors in $K = 2$ networks had a biological significance since there are a small number of cells (attractors) even though the number of genes (N) is large.

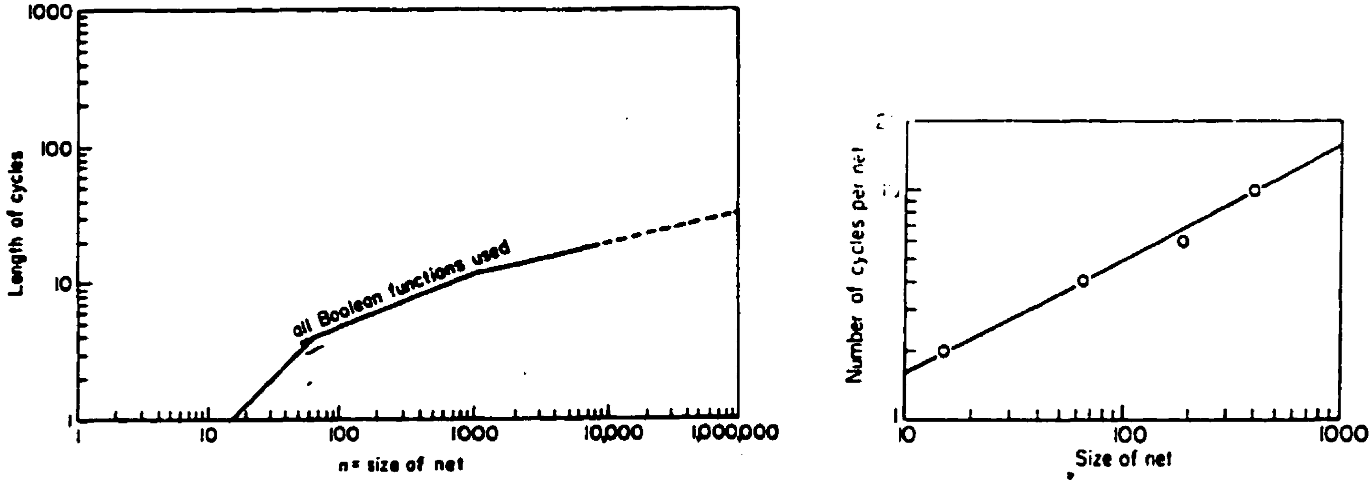


Figure 1.1: Median cycle length and number of cycles as a function of N for networks with $K = 2$. Adapted from [4].

Similar order to that in $K = 2$ networks was found in networks with larger K when certain biases are introduced into the construction of the functions given in the truth tables (Equation 1.1). These biases are effected by the tuning of the p parameter.

To put Kauffman's observations on a firm theoretical foundation, theoretical physicists have focused on: i) demonstrating the existence of a transition in dynamics analogous to order-disorder transitions in statistical physics and identifying the parameter values where the transition occurs [7, 8, 11] ii) studying properties of the model at the biologically relevant 'transition' point. Since the focus of this thesis is on identifying the transition in continuous networks and not specifically on the properties of the network at the transition, I will outline approaches taken on identifying the transition in discrete networks but will only briefly mention the efforts to study properties of discrete networks at the transition.

The Hamming distance H between two different configurations of N elements is defined as the number of positions in the network that contain different values. For example, (000111) and (110011) have a Hamming distance of 3 between them. Derrida and Pomeau analyzed the evolution of the Hamming distance for different values of K and p for a pair of different configurations of the same network [7]. They introduced a variant of the original Kauffman model in which the connections and updating rules of each element are chosen randomly at each time step. This is called the annealed approximation. If the normalized overlap, $\frac{N-H}{N}$, at time t is $a_{12}(t)$, then the fraction of elements that will have all K of their inputs at the same value in both configurations 1 and 2 is, on average, $[a_{12}(t)]^K$. This will force these elements to have identical values in both configurations at the next time step. There is also a contribution to the overlap $a_{12}(t)$ from elements that are in different states between the two configurations. The time evolution of the overlap between two configurations as a function of K and p is [7]:

$$a_{12}(t+1) = [a_{12}(t)]^K + [p^2 + (1-p)^2]\{1 - [a_{12}(t)]^K\}. \tag{1.2}$$

As $t \to \infty$ and $N \to \infty$, this recursion relation has a stable attractive fixed point $a_{12} = 1$ (differing configurations become identical) when $p > p_c$ given by [7]

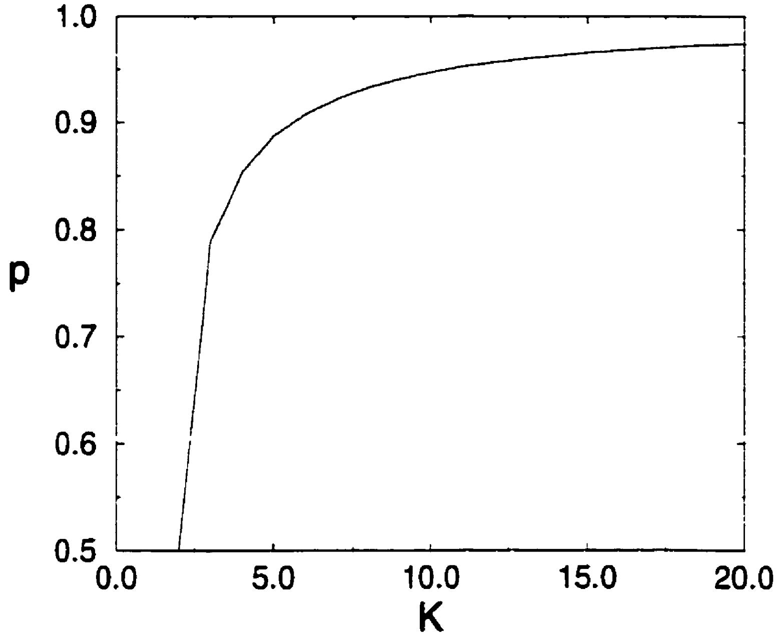


Figure 1.2: Plot of Equation 1.3 in p-K space. This 'critical line' separates ordered from disordered behaviour in the limit that $N \to \infty$.

$$p_c(1 - p_c) = \frac{1}{2K}. \tag{1.3}$$

For p greater than the value given by this expression, the system would typically lie in what is referred to as the 'ordered' regime where steady states exist and the length of cycles $\sim \sqrt{N}$ while for p less than this value the system would be in the 'chaotic' regime where the length of cycles $\sim e^N$. While the term ordered may be appropriate for describing steady state and periodic dynamics, the term chaotic is generally reserved for describing, among other criteria, aperiodic behaviour. Although the improper use of the term 'chaotic' abounds in the literature related to Kauffman networks and cellular automata which both exclude the possibility of aperiodic dynamics, the term chaotic will be used here only to describe dynamics which are aperiodic. The graph of Equation 1.3 in $p - K$ space (shown in Figure 1.2) is called the 'critical line', in analogy with a phase transition at a 'critical point' in the statistical mechanical context. For $p < p_c$ the $a_{12} = 1$ solution becomes unstable and a new

A	B	C	$i(t+1)$
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	0

Table 1.2: If A is frozen to a value of 0, then element i will also be frozen at the next time step.

fixed point emerges with $a_{12} < 1$ such that the Hamming distance between two configurations would approach a non-zero steady state value. In the original Kauffman model the configuration of a system at time t is correlated with the choice of logical functions and inputs for each element. These correlations do not exist in the annealed approximation and thus results from the two models would match only when these correlations become negligible in the original Kauffman model. It is argued in [7] that these correlations vanish in the limit $N \to \infty$.

The concept of a 'stable core' of elements [11] was introduced as an order parameter for random Boolean networks and provided a microscopic explanation of the different behaviours in the network. There is a non-zero probability of having a constant function of all 0's or 1's in the rule table of a given element which decreases rapidly as K is increased. This probability is equal to $p^{2^K} + (1-p)^{2^K}$. These elements become 'frozen' to 0 or 1 after the first time step regardless of their initial values. If one or more frozen elements are inputs to other elements then there is a non-zero probability that these elements will become frozen after the second time step. This process is repeated for subsequent time steps with a larger percentage of elements in the network being frozen and able to propagate their 'freezing effects' to elements that they are inputs to. Flyvberg expresses the fraction of frozen elements as this recursive sum [11]

$$s(t+1)=\sum_{j=0}^{K}\begin{pmatrix}K\\ j\end{pmatrix}s(t)^{K-j}(1-s(t))^{j}p_j. \qquad (1.4)$$

This expression assumes that each element can be an input to another element only once. Therefore this approximation only becomes valid as $N \to \infty$. The steady state solution of this equation is found by setting $s(t+1)=s(t)$ and is displayed in Figure 1.4 for $K=3$ as a function of p [11]. The effect that p has on this quantity and hence on the dynamics is dramatic. Stability analysis of this expression yielded the same critical relation between p and K that was found in the annealed approximation, Equation 1.3.

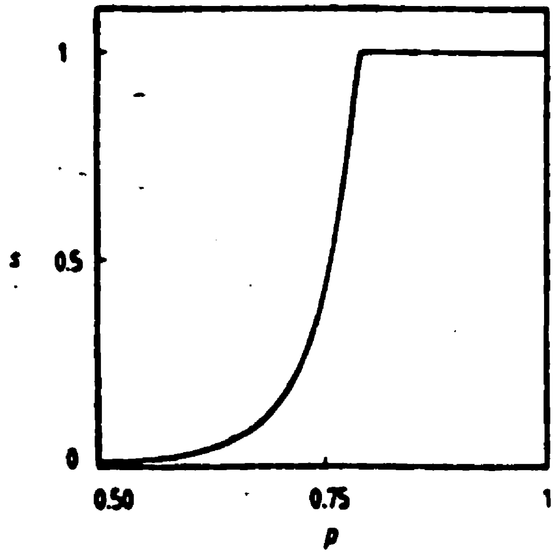


Figure 1.3: Steady state solution of recursion relation for the fraction of frozen elements in a network as a function of p. In this example, $K=3$. Adapted from [14].

More recent works have examined properties of the networks along the 'critical line' Equation 1.3. The scaling behavior of transients and cycles along this 'critical line' at $K=2$ and p=0.5 was examined more carefully [13, 14]. It was found that the power law relationship between N and the median lengths of transients and limit cycles persists over ten orders of magnitude. In the context of the annealed approximation, the closing probabilities for trajectories were studied at $K = 4$, $p = 0.7887$ (from Equation 1.3) [16]. Good agreement was found between the stochastic Markovian model for the evolution of the overlap and the average period of a limit cycle. However, they also found that the distribution of cycle lengths decays much slower in the original model than in the annealed approximation, an effect that becomes more significant with increasing N. This suggests a possible contradiction of the predicted agreement between the annealed and original model as $N \to \infty$ [7].

The critical relation between p and K derived here for discrete networks may have some bearing on characterizing the dynamics in the continuous networks. The derivation in the annealed approximation that led to the expression for criticality depends on the synchronous updating in the Kauffman network and therefore the applicability of this approach to the continuous networks is limited. Also, numerical studies did not investigate the sharpness of this transition as a function of N, K, and p. The concept of the 'frozen core' as a source of order in these networks derives directly from the rule tables for each element and should, in principle, be independent of the updating procedure. The applicability of this approach or similar approaches to the asynchronously updated continuous networks has yet to be investigated.

1.3 Continuous Switching Networks

1.3.1 Definition of Model

In the continuous version of the Kauffman net, the biologically unrealistic constraint of clocked, synchronous updating is dropped and the state of activity of the elements is updated asynchronously: the element that crosses its threshold first is the only one whose Boolean value is changed. The nonlinear sigmoidal response function of a given gene to its inputs is approximated as a step function, resulting in a set of piecewise linear differential equations. As in the discrete case, the model consists of N elements, but here their values are continuous and are represented by a real number, $x_i; i = 1, 2, ..., N$, with a corresponding Boolean variable X_i that is defined as follows [20] :

$$X_i = 0 \text{ when } x_i < 0;$$
$$X_i = 1 \text{ when } x_i \geq 0.$$

The analogous N dimensional differential equation of Equation 1.2 is

$$\frac{dx_i}{dt} = \lambda_i(X_{i_1}(j), X_{i_2}(j), \ldots, X_{i_K}(j)) - \gamma(\vec{X_{(i)}})x_i, i = 1, ..., N. \tag{1.5}$$

where $\lambda_i(X_{i_1}(j), X_{i_2}(j), \ldots, X_{i_K}(j))$ is a scalar whose sign is negative (positive) if the corresponding logical variable $\Lambda_i(X_{i1}(j), X_{i2}(j), \ldots, X_{i_K}(j))$ is 0 (1). The variable $\gamma(\vec{X_{(i)}})$ is without index i because the inhibitory rates are assumed to be the same for each element, and, for simplicity, $\gamma = \gamma(\vec{X}) = 1$. The λ_i are piecewise constant functions whose values may change for particular i when the variables x_i cross their thresholds. The number of elements N correspond to the dimension of phase space. The thresholds divide phase space into 2^N orthants for an N-dimensional system such that the region of phase space $\{x_1 > 0, x_2 < 0, ..., x_N > 0\}$ maps to the Boolean state (10...1), and the the region $\{x_1 > 0, x_2 < 0, ..., x_N < 0\}$ maps to the Boolean state (10...0). The steady state solution of Equation 1.5 within a particular region of phase space, $\frac{dx_i}{dt} = 0; i = 1, ..., N$, is the *focal* point that the system evolves to until crossing into a different orthant of phase space. The focal point is given by $\vec{f}(\vec{X}) = [f_1, ..., f_N]$ and for each element is equal to $f_i(\vec{X_{(i)}}) = \lambda_i(\vec{X_{(i)}})$. It changes when the threshold of an element is crossed, unless the system has evolved to a steady state attractor. The solution of Equation 1.5 can be found through analytic integration. Expressed in terms of $f_i = f_i(\vec{X_{(i)}})$, the solution in region $\vec{X}$ is

$$x_i(t) = f_i + (x_i(0) - f_i)e^t, \tag{1.6}$$

where $x_i(0)$ is the initial value of element i upon entering a given region. When the time evolution of one variable is plotted against the time evolution of another variable, straight lines are produced within each orthant since the exponential time dependence in Equation 1.6 drops out when the slope is calculated. The slope of these straight line trajectories changes discontinuously at each exit wall. The position of a focal point of a region relative to other regions of phase space determine which regions can be entered, and most importantly, which region is entered *first*, or equivalently, which variable crosses its threshold first. A representative phase space trajectory is shown in Figure 1.4 for a 2-dimensional cyclic attractor [25]. The values of each variable at a threshold crossing can be calculated such that trajectories are given an analytic expression. If the j^{th} element crosses its threshold, implying $x_j(t^*)$=0, then the transit time t^* through that region of phase space can be solved by setting Equation 1.6 equal to zero for the j^{th} element. This transit time t^* can then be substituted back into Equation 1.6 for the other variables so that their values at the exit wall of that region when element j crosses its threshold can be expressed as

$$x_i = \frac{x_i(0) - (f_i/f_j)x_j(0)}{1 + x_j(0)/(-f_j)}; \tag{1.7}$$

or, in vector representation, [25],

$$\vec{x^*} = \vec{M}(\vec{x}) = \frac{C\vec{x}(0)}{1 + \vec{c^t}\vec{x}(0)}; \tag{1.8}$$

where $C \in \Re^{n \times n}$. The vector $\vec{c^t} \in \Re^n$ is transposed and has all components equal to zero except for $c_j = -1/f_j$. Equation 1.8 is known as a *fractional linear map* and

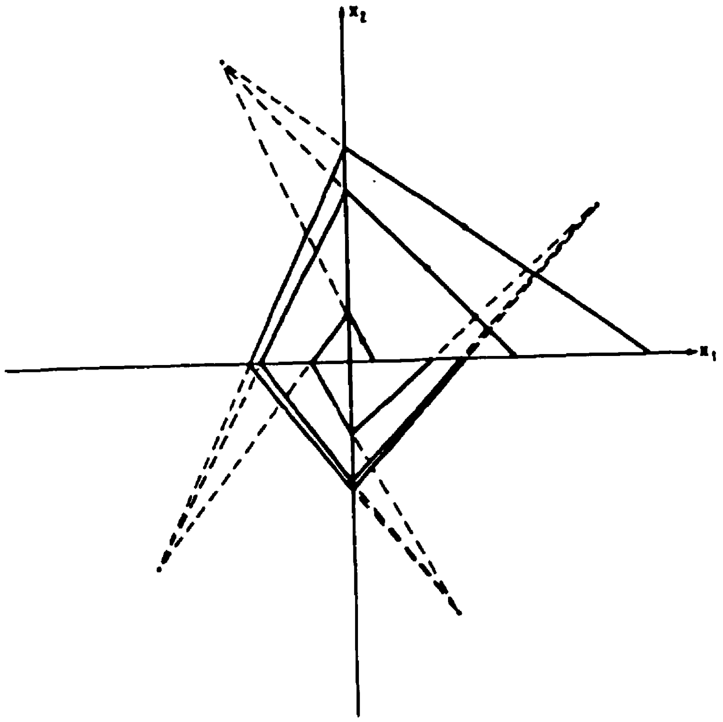


Figure 1.4: Phase space portrait of an $N = 2$ network after a few iterations. This particular flow is that of a cyclic attractor where all trajectories converge to a stable limit cycle. Adapted from [25].

transitions from region to region are computed as a composition of fractional linear maps [23]. The composition of two fractional linear maps remains a fractional linear map and allows significant analysis of the evolution of the system through phase space.

1.3.2 Results

The original work of Glass focused on identifying steady states and limit cycles and predicting their existence based on the logical structure of the network [20]. This was aided through the mapping of the Boolean configuration of the system, $\vec{X}(t)$, onto the vertices of an N-dimensional hypercube; and the evolution of the system onto flows along the edges of this object. When elements are restricted from being inputs to themselves, flows in phase space can be represented on a hypercube in which each edge is directed in only one of the two possible directions. A flow along the edges of a hypercube for a particular 4-dimensional system is shown in Figure 1.5 [23].

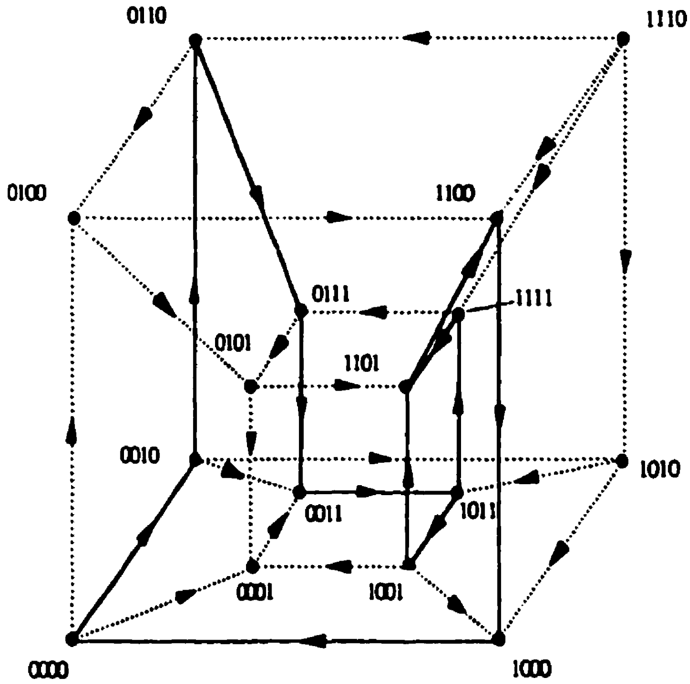


Figure 1.5: Hypercube mapping of the flow of a trajectory for $K = 3$. Edges with bold face arrows represent allowed transition from one state to an adjacent state. Adapted from [23].

Glass defined a hypercube with a particular orientation of directed edges as a *state transition diagram*. Two systems whose state transition diagrams can be superimposed upon each other under a symmetry operation of the N-cube (e.g. rotation or reflection about a particular axis or vertex of the hypercube) are said to belong to the same *equivalence class*. Glass studied properties of these equivalence classes to connect the properties of various steady states and limit cycles with the geometric structure of flows on the hypercube. In particular, it was shown for $N \geq 3$ that a cyclic attractor on the hypercube implies a unique stable limit cycle in the associated piecewise linear differential equations when the focal points are ± 1 [25].

Studying these piecewise linear equations in the context of neural networks, Lewis and Glass found numerical evidence for the existence of a chaotic attractor in a 6-dimensional network [21, 22]. A 5-dimensional Poincaré section that the solution trajectory of their system crossed repeatedly was projected on to a particular 2-dimensional plane. In this way, a density histogram of the values of a particular element was plotted for long times. Neighboring trajectories diverged over time and the corresponding density histograms were invariant with successive iterations. This represented numerical evidence for the existence of two features commonly found in many chaotic systems: ergodicity and a unique invariant density [21].

A particular four dimensional network was discovered, from a search of many thousands, that displayed chaotic dynamics. Mestl [23] demonstrated that it is not possible to have chaos in 3-dimensions, implying that this example existed in the lowest dimension possible. During iteration, the dynamics follows the solid edges in Figure 1.5.

A Poincaré section was chosen at the 'wall' separating vertices (0011) and 1011. Numerical studies demonstrated that the dynamics could be described by a *strange attractor* (i.e. fractal structure).

Further analysis was carried out for evaluating the Lyapunov exponent which is a measure of the divergence of trajectories in phase space. As described in Section 1.3.1, the trajectory of the system through phase space can be analytically expressed as the composition of fractional linear maps. The Jacobian of this fractional linear map after one iteration is

$$J^1 = \frac{\partial M}{\partial \vec{x}} |\vec{x^0} \tag{1.9}$$

where $M = \vec{M}(\vec{x})$ is as defined in Equation 1.8. With

$$J^n = \frac{\partial M}{\partial \vec{x}} |\vec{x^{n-1}} \dots \frac{\partial M}{\partial \vec{x}} |\vec{x^0} \tag{1.10}$$

representing the product of the Jacobians after n iterations of the fractional linear map. With $\vec{u^0}$ representing an initial perturbation, the Lyapunov exponents are expressed as

$$h = h(\vec{x^0}, \vec{u^0}) = \lim_{n \to \infty} \frac{1}{n} \ln(|J^n \cdot \vec{u}^{(0)}|). \tag{1.11}$$

There are at most N distinct Lyapunov exponents, where N is the dimension of the system. Using analytical methods described in [23], the lower bound on the largest Lyapunov exponent in this system is found to be 0.31, verifying the numerical evidence for sensitive dependence on initial conditions and chaotic dynamics. The Lyapunov exponent, related to the rate of convergence/divergence of nearby trajectories, was calculated using this technique. It was approximately equal to 0.45, indicating chaotic dynamics since it was positive. Mestl et al found that chaotic attractors are rare in low dimensions, but ubiquitous in high-dimensional networks with a moderate number of inputs (~ 10) [24].

While the existence of steady states, limit cycles, and deterministic chaos has been demonstrated in these continuous networks, there has been little investigation into the mechanisms which determine which kind of dynamics predominate. There has been analytical studies of the transition from steady state to limit cycle dynamics as a function of the biasing parameter p in the discrete networks, but there has not been a systematic analysis of the dynamics as a function of p in the continuous networks. This work is an attempt to fill this void.

2

Ordered and Disordered Dynamics in Random Networks

Genetic networks have been modeled by random Boolean networks in which time is discrete and each element computes a Boolean function based on the values of inputs to that element [3]. Since the number of human genes is of the order of 100,000, and each gene is idealized as either on (1) or off (0), the state space for the human gene activity is huge. An order-disorder transition has been described for random Boolean networks in the limit that the number of variables, $N \to \infty$, as a function of the number of inputs per variable, K, and the probability, p, that the truth table for a given element will have a bias for being 1 [7, 10, 11, 13, 15, 16, 26]. The order-disorder boundary is given by

$$K_c = \frac{1}{2p_c(1-p_c)}, \tag{2.1}$$

where K_c and p_c represent the values of K and p on the boundary [7, 11, 16]. Kauffman has argued that for a network to be biologically meaningful, it should have relatively few attractors, and the cycle length of attractors should be comparatively short [3]. In real biological systems there are not clocking devices to generate synchronous updating and theoretical models are more appropriately formulated as continuous differential equations [20, 23, 24, 27, 28, 29, 30]. Here we present numerical evidence for an order-disorder transition in differential equation analogues of the discrete switching networks. We also present a probabilistic model of the dynamics, in which we show that (2.1) also applies to the continuous equations.

First consider a logical network consisting of N binary variables, $X_i = 0, 1,\ i = 1, \ldots, N$. Since we consider the networks as models of genes, X_i represents the activity of gene i. In other contexts, logical variables may represents spins or voltages. The network is updated by means of the dynamical equation

$$X_i(j+1) = \Lambda_i(X_{i_1}(j), X_{i_2}(j), \ldots, X_{i_K}(j)), \qquad i = 1, \ldots, N, \tag{2.2}$$

where $\Lambda_i(X_{i_1}(j), X_{i_2}(j), \ldots, X_{i_K}(j)) \in \{0,1\}$ and K is the number of inputs. More compactly, we have

$$\mathbf{X}(j+1) = \Lambda(\mathbf{X}(j)), \qquad i = 1, \ldots, N, \tag{2.3}$$

Thus, for any state $\mathbf{X}(j)$, Λ is a truth table determining $\mathbf{X}(j+1)$.

The logical structure of Eq. (2.2) can be captured by a differential equation [20]. To a continuous variable $x_i(t)$, we associate a discrete variable $X_i(t)$,

$$X_i(t) = 0 \text{ if } x_i(t) < 0; \text{ otherwise } X_i(t) = 1. \tag{2.4}$$

For any logical network, we define an analogous differential equation,

$$\frac{dx_i}{dt} = -x_i + \lambda_i(X_{i_1}(j), X_{i_2}(j), \ldots, X_{i_K}(j)), \qquad i = 1, \ldots, N, \tag{2.5}$$

where $\lambda_i(X_{i_1}(j), X_{i_2}(j), \ldots, X_{i_K}(j))$ is a scalar whose sign is negative (positive) if the corresponding logical variable $\Lambda_i(X_{i1}(j), X_{i2}(j), \ldots, X_{i_K}(j))$ is 0 (1).

For each variable, the temporal evolution is governed by a first order piecewise linear differential equation. Let $\{t_1, t_2, \ldots, t_k\}$, denote the *switch times* when any variable of the network crosses 0. The solution of Eq. (2.5) for each variable x_i for $t_j < t < t_{j+1}$, is

$$x_i(t) = x_i(t_j)\, e^{-(t-t_j)} + \lambda_i(X_{i_1}(j), X_{i_2}(j), \ldots, X_{i_K}(j))(1 - e^{-(t-t_j)}). \tag{2.6}$$

Depending on the particular network, Eq. (2.5) can display steady states, limit cycles, quasiperiodicity or chaos [20, 23, 27, 28, 29, 30]. Although the origin of chaos in one particular four-dimensional network has been analyzed [23], no general methods have yet been developed to determine whether chaotic dynamics exists in any given equation without integrating it. Chaos is the usual behavior in these networks for $N = 64$, $9 \le K \le 25$, $p = 0.5$ [24].

In carrying out the computations, we have made several additional assumptions concerning the structure of the equations. (i) There is no self-input or reciprocal input. This means that i cannot be an input to itself, and if i is an input to k, k is not also an input to i. These assumptions eliminate the possibility of asymptotic approach to stable steady states in the neighborhood of threshold axes $x_i = 0$ [20, 29, 30]. If these assumptions are not made, the techniques for analysis of the dynamics based on the symbolic dynamics described below fail. (ii) We set λ_i in Eq. (2.5) to be in the range $-1\pm \mid .01 \mid$ or $1\pm \mid .01 \mid$. The focal points of individual elements are given a small perturbaton away from 1 or -1 that is some fraction of the initial value of each element. This ensures that the situation where two variables cross their threshold at the same time does not arise. To determine the effect that this arbitrary perturbation has on the classification of the attractors of networks, the strength of this perturbation is varied among -0.0001,-0.001,-0.01,-0.1,+0.1, +0.01,+0.001,+0.0001 of the initial value of each element for $K = 6$, $N = 50$, and $p = 0.75$. The results indicate no systematic variation above the expected statistical noise level of $\frac{1}{\sqrt{no.of nets.}}$ or 10% for this case of 100 networks. (iii) In the truth tables we assume that for a given value of p, half the variables of the network have entries that are biased towards 1 and half are biased towards 0.

Although the piecewise-linear nature of Eq. (2.5) facilitates the speed and accuracy of numerical integration over other integration techniques, numerical studies nevertheless grow large rapidly. Consequently, we focus on the order-disorder transition for limited regions of parameter space. The system was simulated using a source code written in FORTRAN. Equations 2.5 were integrated from random initial conditions. The random number generator employed was obtained from Numerical Recipes. The rule tables and the connection topology are randomly constructed for the specified values of N, K, and p. The code was compiled and run on a PC with an Intel Pentium processor. The operating system used was Redhat Linux and the version of FORTRAN used was FORTRAN 77. For each network, we investigated one initial condition and classified the dynamics after a transient of 120,000 switch times. The typical length of time to evolve the system for this many time steps is a function of N, K, and p. For $N \sim 100$, $K \sim 10$, and $p = 0.5$, it would take $\sim$ 5 hours to run 50 networks for 120,000 switch times. The dynamics can be classified by using symbolic dynamics, keeping track of the variable that switches as a function of time. Provided the network does not reach a steady state during the course of the integration, we generate a sequence of integers denoting the label of the variable that switches at each consecutive switch time. We search for periodicities in this sequence of integers, searching for periodicities up to length 3000. When a periodic sequence is identified, additional sequences of integers are examined to verify periodicity.

Given the restrictions on the equations mentioned above, if the symbolic sequence is periodic, in the differential equation there is a stable limit cycle oscillation [20]. If the sequence is not periodic, then in the differential equation, there can either be quasiperiodic dynamics or chaotic dynamics. To identify quasiperiodic dynamics, we need to keep track of the exact switching times of all the variables of the network. However, pilot investigations of several hundred randomly constructed networks failed to identify quasiperiodic dynamics. Consequently, to simplify the analysis we lump quasiperiodic rhythms together with chaotic networks, where we expect that the incidence of quasiperiodic dynamics is negligible (less than 1%).

Figure 2.1 illustrates typical dynamics for a single variable in a network with $N = 50$, $K = 8$. The same connection matrix was assumed for p incremented in from 0.95 to 0.50 in steps of 0.05, although only three values, $p = 0.5, 0.8, 0.85$ are displayed. When $p = 0.75$ the network evolves to a steady state (not shown). In Fig. 2.1(a) and 2.1(b), the symbolic transition sequences corresponding to the time series do not show periodicities. The Lyapunov number, which can be evaluated numerically using techniques in Ott [24, 31], is positive for panels (a) and (b) and negative for panel (c). Based on the above, the dynamics are deterministic chaos in panels (a) and (b) and periodic in panel (c). The length of the limit cycle displayed in (c) is 24 threshold crossings and the period is 3.744 time units. Based on the above data it follows that lowering p does not always cause a single network to tend toward chaotic behaviour. This network went from periodic dynamics to a steady state to deterministic chaos as p was lowered.

Figure 2.2 shows the numbers of networks displaying steady states, limit cycles, and chaos for $K = 8$ and $N = 50$ panel (a), $N = 100$ panel (b), and $N = 200$ panel (c), for 50 different randomly generated networks for $p \in [0.5, 1.0]$ incremented in steps of 0.02. As p increases the number of chaotic networks decreases and the number of networks displaying steady states increases. The values of p associated with limit cycles are centered in an increasingly narrow range as N increases. The width-half-max of the limit cycle region is designated by w, and the midpoint of the limit cycle region is considered to be the critical value of p for that particular value of N. With additional system sizes simulated, a scaling relation of the form $w \sim N^{-\delta}$ could be found for the width-half-max. Extrapolating to $N \rightarrow \infty$ we would find that the width-half-max goes to zero. Since the total number of orthants of phase space that are visited during the simulations is a tiny fraction of the 2^N orthants of phase space, we cannot exclude the possibility that networks classified as chaotic will eventually reach a stable steady state or limit cycle if the integration times are extended.

Figure 2.3 shows the number of networks displaying deterministic chaos for $N = 50$ and $K = 6$ panel (a), $K = 8$ panel (b) when 320,000 switch times are followed before classifying the dynamics. For a given value of p the fraction of chaotic networks is higher for $K = 8$ than for $K = 6$. The number of chaotic networks approaches 0 (corresponding to the critical value of p) at a lower value of p for $K = 6$ than for $K = 8$.

Figure 2.4 illustrates the variation in the numbers of steady states, limit cycles, and chaotic dynamics for $N = 50$, $K = 6$, as a function of the length of the transient. As the transient length is increased, the number of networks that pass the criterion for chaotic dynamics decreases. For higher values of p almost all the networks are in steady states or cycles and the length of the transient before this behavior is established is short. However, for $p = 0.5$, increasing the transient from 80,000 to

320,000 transitions led to changes in identification of chaotic networks – the fraction of chaotic attractors decreased from .90 to .80. These results illustrate the problems of numerical analysis. Even for the comparatively small $N = 50$ networks, the state space is huge, and there may be weak attraction to sparsely scattered steady states or limit cycles. Thus, these methods cannot exclude the possibility that for any network classified as chaotic, an attractor would be found had the numerical simulation been arbitrarily long.

We now characterize the dynamics in Eq. (2.5). Let j designate the time interval (t_j, t_{j+1}), in Eq. (2.6). At any given time in the interval j, the state in the continuous equation is mapped to the logical state $\mathbf{X}(j)$. The distance between two logical states is the number of variables in which the activities differ. We call the distance between $\mathbf{X}(j)$ and $\mathbf{X}(j+1)$, determined from Eq. (2.2), the *outflow dimension*, $h(j)$, of $\mathbf{X}(j)$. The outflow dimension is a measure of the number of different variables that have the potential to cross 0 at the next switch time of the network [20]. If $h(j) = 0$, then the system will approach a steady state.

Using methods similar to those developed in earlier work by Derrida and others [7, 26, 15], we compute the mean value of h. Denote the number of variables in logical state α, and with truth table entry β in time interval j by $N_{\alpha\beta}(j)$, where $\alpha, \beta \in \{0, 1\}$. If $N_{01}(j) = N_{10}(j) = 0$, there is a fixed point. If this is not the case there will eventually be a transition to a new state. The probability that there will be a transition of a variable from state 0 to state 1 is $P_{0\to1} = N_{01}(j)/(N_{01}(j) + N_{10}(j))$, and the probability there will be a transition of a variable from state 1 to state 0 is $P_{1\to0} = N_{10}(j)/(N_{01}(j) + N_{10}(j))$.

The expected number of inputs from the switching variable to variables in state (α, β) is $\rho N_{\alpha\beta}$, where $\rho = K/N$. Now assume that there is a transition of a variable from $0 \to 1$ at switch time t_{j+1}. We adopt a probabilistic approach to determine $N_{\alpha\beta}(j+1)$. Consider first the value of $N_{00}(j+1)$. This may be different from $N_{00}(j)$ if inputs from the variable that changed its state have inputs to variables in N_{00} that leads to a change to N_{01}, or inputs to variables in N_{01} that leads to changes to state

N_{00}. Thus, we find that,

$$N_{00}(j+1) = N_{00}(j) - \rho(1-p)N_{00}(j) + \rho p N_{01}(j).$$

If there is a transition of a variable from 1 → 0 at time t_j, then the above expression would be changed to

$$N_{00}(j+1) = N_{00}(j) + 1 - \rho(1-p)N_{00}(j) + \rho p N_{01}(j).$$

In similar fashion, by weighting the respective transition probabilities, we are led to the following system of equations for $N_{\alpha\beta}(j+1)$.

$$\begin{aligned} N_{00}(j+1) &= N_{00}(j) + P_{1\to 0} - \rho(1-p)N_{00}(j) + \rho p N_{01}(j) \\ N_{01}(j+1) &= N_{01}(j) - P_{0\to 1} + \rho(1-p)N_{00}(j) - \rho p N_{01}(j) \\ N_{10}(j+1) &= N_{10}(j) - P_{1\to 0} - \rho(1-p)N_{10}(j) + \rho p N_{11}(j) \\ N_{11}(j+1) &= N_{11}(j) + P_{0\to 1} + \rho(1-p)N_{10}(j) - \rho p N_{11}(j) \end{aligned} \tag{2.7}$$

At steady state, we have $N_{\alpha\beta}(j+1) = N_{\alpha\beta}(j) = N^*_{\alpha\beta}$, where the asterisk represents the steady state value. Substituting this relation in Eq. (2.7), we find that $N^*_{01} = N^*_{10}$. Substituting this result in the Eq. (2.7), we find

$$\begin{aligned} -\frac{1}{2} &= -\rho(1-p)N^*_{00} + \rho p N^*_{10}, \\ -\frac{1}{2} &= \rho(1-p)N^*_{10} - \rho p N^*_{11}. \end{aligned} \tag{2.8}$$

Using the conservation condition, $N^*_{00} + 2N^*_{10} + N^*_{11} = N$, we have three simultaneous equations. Solving these equations and recalling that at steady state, the mean outflow dimension, h^* is given by $N^*_{10} + N^*_{01} = 2N^*_{10}$, we compute

$$\frac{h^*}{N} = \frac{1}{K}(-1 + 2Kp - 2Kp^2). \tag{2.9}$$

In Fig. 2.5(a), we compare the theoretical estimates from Eq. (2.9) (solid curve) with the numerical computations for $N = 50$, 100, 200. The theoretical estimate lies consistently above the numerically computed values, but shows a similar functional dependence on p. When $h^* = 0$, we recover Eq. (2.1). In Fig. 2.5(b), we plot the fraction of chaotic networks as a function of h^*. The transition occurs approximately in the range $0 < h^* < 20$ for all values of N considered. The results in Figure 2.5 provide a challenge for further theoretical analysis.

These results have connections with the extensive studies carried out on the discrete time and discrete state space switching networks [7, 10, 11, 26, 16, 13, 15]. In contrast to the earlier work, in which all finite networks must eventually cycle in the limit $t \to \infty$ and which do not therefore admit deterministic chaos, in Eq. (2.5) deterministic chaos is possible [23. 24]. At the moment. there are no general techniques to assert deterministic chaos in any given network, and it is possible that eventually networks identified as chaotic here will reach limit cycles or steady states. Nevertheless, the extremely long transient behavior would appear to render these architectures improbable for the highly constrained dynamics in real biological systems. In the continuous equations, the critical line (2.1) defines the line at which $h^* = 0$ and almost all networks with $p \geq p_c$ display steady states. Thus, the continuous equations show transitions in dynamics even for finite N. As the value of p is varied from 0.5, there is a transition from chaotic dynamics to steady states, with an intervening zone of periodic dynamics that becomes increasingly narrow as N increases.

Figures

1. Time series for a typical variable in a network with $N = 50$, $K = 8$. (a) $p = 0.5$, (b) $p = 0.8$, (c) $p = 0.85$. In panels (a) and (b), the time series satisfy our tests for chaos and in (c) the dynamics are periodic.

2. Variation in the number of networks displaying steady states (circles), limit cycles (squares), and chaos (diamonds) for $K = 8$, as a function of p, (a) $N = 50$, (b) $N = 100$, (c) $N = 200$. A single initial condition was selected for each of 50 different networks for each value of p. Networks were iterated for 120,000 switch times.

3. Fraction of chaotic networks for 320,000 iterations as a function of p for $N = 50$ when (a) $K = 6$, (b) $K = 8$. The fraction of chaotic networks approaches zero at a greater value of p for $K = 8$ than for $K = 6$.

4. Variation in the numbers of chaotic attractors for $N = 50$, $K = 6$, as a function of the length of the transient. The circles designate a transient length of 10,000 iterations; squares correspond to 80,000 iterations; and diamonds correspond to 320,000 iterations. As the transient length is increased, the number of networks that pass the criterion for chaotic dynamics decreases.

5. (a) Normalized mean outflow dimension, h^*/N, as a function of p for $N = 50$ (triangles), $N = 100$ (asterisks), $N = 200$ (plus signs) compared with the theoretical result from Eq. (2.9). For each network, the value of h was averaged over 3000 switch times following a transient of 120,000 switch times. (b) Fraction of chaotic networks as a function of h^* for $N = 50$ (triangles), $N = 100$ (asterisks), $N = 200$ (plus signs).

Figure 2.1

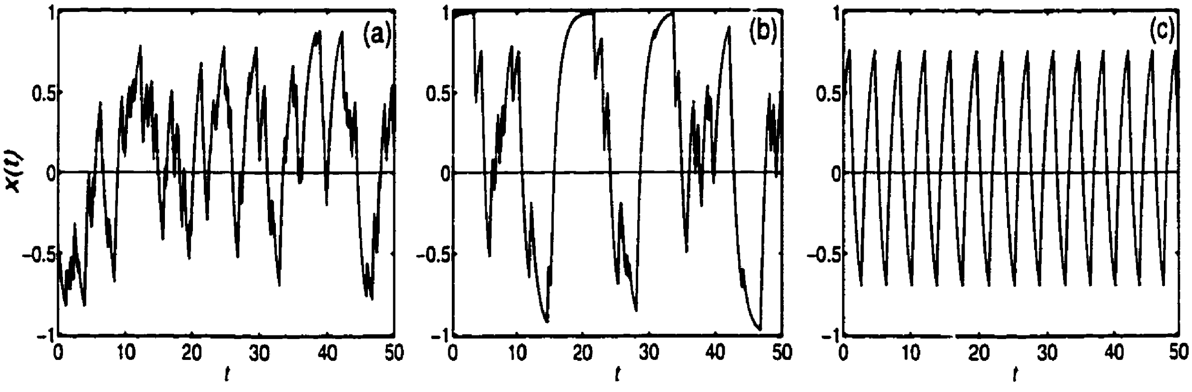


Figure 2.2

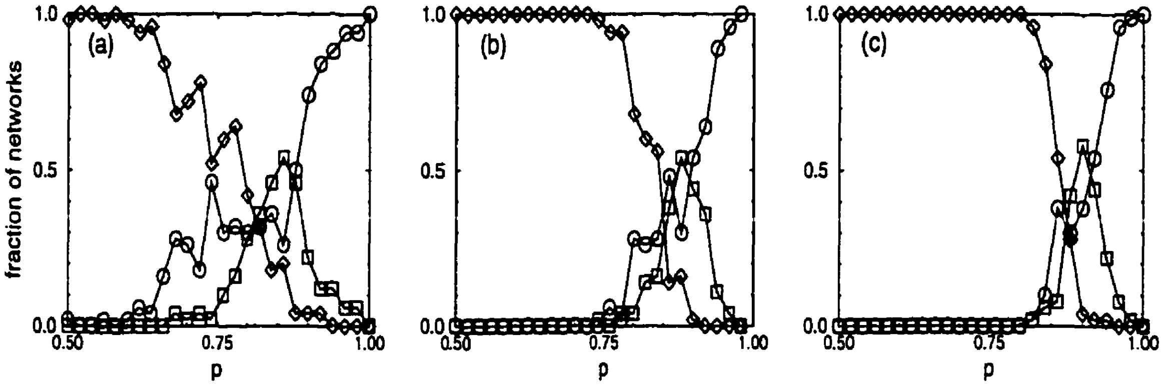


Figure 2.3

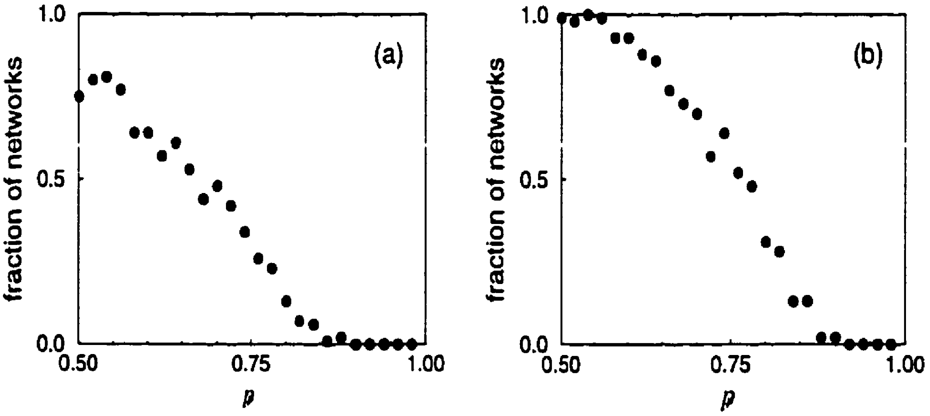

1.0
0.5
0.0
fraction of networks
(a)
0.50
0.75
1.00
p
1.0
0.5
0.0
fraction of networks
(b)
0.50
0.75
1.00
p

Figure 2.4

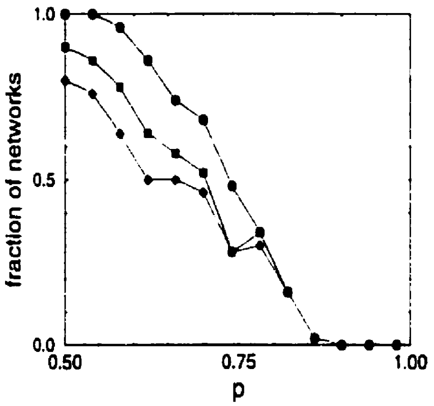
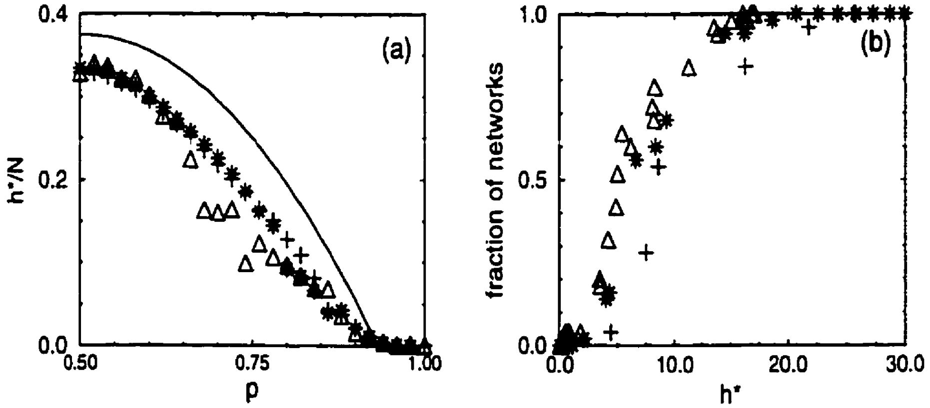


Figure 2.5


0.4
0.2
0.0
h*/N
0.50
0.75
1.00
p
(a)
1.0
0.5
0.0
fraction of networks
0.0
10.0
20.0
30.0
h*
(b)

3

Discussion and Conclusion

The order-disorder transition demonstrated in discrete logical switching networks as the biasing parameter p is tuned is also found in continuous analogues of the discrete networks. Simulation results for $K \geq 6$ and $N \geq 50$ show that when $p \approx 1.0$, all of the randomly selected networks reach a steady state. As p is lowered, the number of steady state attractors decrease and the number of periodic and chaotic attractors increase. When p approaches 0.5, the fraction of chaotic attractors approaches 1. When N is increased, the range of p over which the transition from steady state to chaos occurs becomes narrower. Periodic dynamics only exist in this narrowing transition region. I speculate that the width of this transition region goes to zero as $N \to \infty$. Thus, in the current context, the felicitous phrase "edge of chaos", which is sometimes used to characterize parameter values near the critical line, refers to a set of parameter values where stable limit cycle oscillations predominate. The value of p where this transition occurs increases as K increases.

A probabilistic model for the dynamics was used to predict qualitative features of the system and to identify the transition point. The average number of elements that can cross their threshold at a given switch time, h^*, is the mean dimension of the subspace admitting outflows during the integration of the equations. This quantity was computed from a Markovian model for the dynamics. Changes in h^* correlate with changes in the proportion of steady state, periodic, and chaotic attractors. When the expression for h^* is set equal to zero, indicating the onset of steady state dynamics, a relation between p and K corresponding to this transition is found. The critical relation between p and K is identical to Equation 1.3, the boundary in $p-K$ space between ordered and disordered dynamics in discrete switching networks.

One aspect that requires further analysis is the effects of iteration time. It shows that when p is near 0.5, the number of networks that reach steady state or periodic attractors increases as the number of iterations is increased. The fraction of chaotic networks as the length of the transient is increased does not approach a constant value and therefore it is not possible to estimate a lower limit on the fraction of chaotic networks. Numerical investigations of this issue of convergence of the number of chaotic attractors requires a great amount of computer time and therefore analytical methods need to be employed to solve this problem.

If the state of a gene is determined by the value of just one of its inputs, regardless of the values of the other inputs, this Boolean function is said to be *canalizing*. These functions can serve the biologically important role of making the state of a gene robust against changes or perturbations of the non-canalizing input genes. Studies of eukaryotic gene regulation suggest that there exists a bias in gene interactions toward canalizing functions [32]. The percentage of canalizing functions dramatically affect the dynamics of discrete switching networks [4, 5]. I have recently demonstrated a transition from ordered to chaotic dynamics in continuous networks, similar to the transition that occurs as p is varied, as the percentage of canalizing inputs is lowered.

The late physicist John Von Neuman stated that the understanding of complex systems composed of many interacting parts would advance from studying the detailed mechanisms and interactions between elementary units and also the integration of the interacting components into a functioning whole. The merging of insights from these two approaches may soon provide a deep and comprehensive understanding of how gene networks really work. It has recently become possible to make large scale measurements of the level of gene expression over time such that predictions from theoretical models can be compared to experiments [33]. From this time series data it would be possible to infer the structure of the gene network [34]. Kauffman's assumption of a randomly constructed network was made in the face of the overwhelming complexity of unknown gene interactions. It is now becoming possible to unravel the exact nature of interactions between genes. Knowledge of the structure of gene networks, whether inferred or found by direct observation, could have great impact on the understanding of embryonic development.

Whether in the study of model gene or neural networks, spin systems, or turbulent flow, the problem central to the study of complex dynamical systems composed of many interacting components is this: given the properties of the microscopic components and their interactions, predict the macroscopic global dynamics of the system. While this task has been accomplished for systems that fall under the purview of equilibrium statistical mechanics, for most systems of interest that pervade our non-equilibrium world, few inroads have been made. In the context of continuous gene networks, the little progress made on this front has been restricted to low dimensional networks [20, 35, 23]. No techniques exist to predict the dynamics of a high dimensional network based on its logical structure and connective topology. Insights into the relation between the structure of these networks and their dyanamics will ultimately provide a deep understanding into the fundamental mechanisms responsible for the development and maintenance of living organisms.

BIBLIOGRAPHY

[1] W.S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, **5**:115, 1943.

[2] J.D. Cowan. The problem of organismic reliability. *Progress in Brain Research: Cybernetics of the Nervous System (N. Wiener and J.P. Schade, eds).*, **17**:9, 1965.

[3] S.A. Kauffman. Metabolic stability and epigenesis in randomly connected nets. *J. Theoret. Biol.*, **22**:437, 1969.

[4] S.A. Kauffman. Gene regulation networks: A theory for their global structure and behavior. *Current Topics in Dev. Biol.*, **6**:145, 1971.

[5] S.A. Kauffman. The large scale structure and dynamics of gene control circuits: An ensemble approach. *J. Theoret. Biol.*, **44**:167, 1974.

[6] F. Fogelman-Soulie. Frustration and stability in random boolean networks. *Disc. Appl. Math.*, **9**:139, 1984.

[7] B. Derrida and Y. Pomeau. Random networks of automata: A simple annealed approximation. *Europhys. Lett.*, **1**:45, 1986.

[8] B. Derrida and G. Weisbuch. Evolution of overlaps between configurations in random boolean networks. *J. Physique*, **47**:1297, 1986.

[9] H.J. Hilhorst and M. Nijmeijer. On the approach of the stationary state in Kauffman's random boolean network. *J. Physique*, **48**:185, 1987.

[10] B. Derrida and H. Flyvberg. The random map model: a disordered model with deterministic dynamics. *J. Physique*, **48**:971, 1987.

[11] H. Flyvberg. An order parameter for networks of automata. *J. Phys. A: Math. Gen.*, **21**:L955, 1988.

[12] H. Flyvberg and N.J. Kjaer. Exact solution of Kauffman's model with connectivity one. *J. Phys. A: Math. Gen.*, **21**:1695, 1988.

[13] A. Bhattacharijya and S. Liang. Power-law distribution in some random boolean networks. *Physical Review Letters*, **77**:1644, 1995.

[14] A. Bhattacharijya and S. Liang. Median attractor and transients in random boolean nets. *Physica D*, **95**:29, 1996.

[15] B. Luque and R. Sole. Phase transitions in random networks: Simple analytic determination of critical points. *Physical Review E*, **55**:260, 1997.

[16] U. Bastolla and G. Parisi. Closing probabilities in the Kauffman model: An annealed computation. *Physica D*, **98**:1, 1996.

[17] U. Bastolla and G. Parisi. The critical line in Kauffman networks: A numerical study. *J. of Theor. Biol.*, **187**:117, 1997.

[18] U. Bastolla and G. Parisi. Relevant elements, magnetisation and dynamical properties in Kauffman networks: A numerical study. *Physica D*, **115**:203, 1998.

[19] U. Bastolla and G. Parisi. The modular structure of Kauffman networks. *Physica D*, **115**:219, 1998.

[20] L. Glass. Classification of biological networks by their qualitative dynamics. *J. theor. Biol.*, **54**:85, 1975.

[21] J. Lewis and L. Glass. Steady states, limit cycles, and chaos in models of complex biological networks. *International Journal of Bifurcation and Chaos*, **1**:477, 1991.

[22] J. Lewis and L. Glass. Nonlinear dynamics and symbolic dynamics of neural networks. *Neural Computation*, **4**:621, 1992.

[23] T. Mestl, C. Lemay, and L. Glass. Chaos in high-dimensional neural and gene networks. *Physica D*, **98**:33, 1996.

[24] T. Mestl, R.J. Bagley, and L. Glass. Common chaos in arbitrarily complex feedback networks. *Physical Review Letters*, **79**:653, 1997.

[25] L. Glass and J. Pasternack. Stable oscillations on mathematical models of biological control systems. *J. theor. Biol.*, **6**:207, 1978.

[26] G. Weisbuch. *Complex System Dynamics.* Addison-Wesley, Redwood City, 1991.

[27] R. Thomas and R.D. Àri. *Biological Feedback.* CRC Press, Boca Raton, 1990.

[28] R.J. Bagley and L. Glass. Counting and classifying attractors in high dimensional dynamical systems. *J. Theor. Biol.*, **183**:269, 1996.

[29] T. Mestl, E. Plahte, and S.W. Omholt. *J. Theor. Biol.*, **176**:291, 1995.

[30] T. Mestl, E. Plahte, and S.W. Omholt. *Dyn. Stab. Systems*, **2**:179, 1995.

[31] E. Ott. *Chaos in Dynamical Systems.* Cambridge University Press, Cambridge, 1993.

[32] S. Harris, B. Sawhill, A. Wuensche, and S. Kauffman. Biased eukaryotic gene regulation rules suggest genome behavior is near edge of chaos. *Santa Fe preprint*, 1997.

[33] S. Fuhrman, X. Wen, G. Michaels, and R. Somogyi. Genetic network inference. *Proc. of the International Conf. on Complex Systems*, 1997 (in press).

[34] S. Liang, S. Fuhrman, and R. Somogyi. Reveal, a general reverse engineering algorithm for inference of genetic network architectures. *Proc. Pacific Symposium on Biocomputing*, 1998 (in press).

[35] L. Glass. *Combinatorial aspects of dynamics in biological systems. In Statistical Mech. and Stat. Methods in Theory and Application.* (U. Landman, ed.). Plenum, New York, 1977.

[36] J.J. Hopfield. Neural networks and physical systems with emerging collective computational ability. *Proc. Natl. Acad. Sci. USA*, **79**:2554, 1982a.

[37] F. Jacob and J. Monod. *Genetic repression, allosteric inhibition and cellular differentiation. In Cytodifferentiation and Macromolecular Synthesis (M. Locke, ed).* Academic Press, New York, 1963.

[38] S.A. Kauffman. *Origins of Order.* Oxford University Press, Oxford, 1993.

[39] S. Wolfram. *Cellular Automata and Complexity: Collected Papers.* Addison-Wesley, Reading, MA, 1994.

[40] B. Sawhill and S. Kauffman. Phase transitions in logic networks. *Santa Fe preprint*, 1996.

[41] L. Glass and C. Hill. Ordered and disordered dynamics in random networks. *Europhysics Letters*, **41**:599, 1998.