Extreme Chaos in Boolean Networks

Winfried Just[1]
German A. Enciso[2]

1. Department Mathematics, Ohio University

2. Mathematical Biosciences Institute and Harvard Medical School

This material is based upon work supported by the NSF under Agreement No. 0112050 and by The Ohio State University.

October 27, 2008

Boolean networks

Let $\Pi = \{0, 1\}^n$, where $n \geq 1$. A map $g : \Pi \to \Pi$ defines an n-**dimensional Boolean network** or **Boolean system** (Π, g).

A state in the system at time t will be denoted by $s(t) = [s_1(t), \ldots, s_n(t)]$.

The dynamics is given by

$$s(t+1) = g(s(t)), \quad s(t) \in \Pi.$$

Note that all trajectories eventually reach a periodic orbit or a fixed point.

The letter n will always stand for the dimension of the system.

Models of gene regulation

Genes code *gene products* (usually proteins). When a gene is *expressed*, its product gets manufactured. After some time delay, the presence/absence of gene products can switch expression of certain genes on or off.

This can be crudely modeled by Boolean networks if we interpret a value $s_i(t) = 1$ as gene number i being expressed at time t and a value $s_i(t) = 0$ as gene number i not being expressed at time t.

It has been argued that Boolean models of gene regulation should show mostly *ordered* dynamics.

Chaotic vs. ordered dynamics

Empirical studies show that Boolean networks dynamics falls either into the *ordered regime* with relatively *short attractors*, a large proportion of *eventually frozen nodes*, and high *homeostatic stability;* or into the *chaotic regime* with relatively *long attractors*, very few *eventually frozen nodes*, and low *homeostatic stability.*

Empirical studies of random Boolean networks indicate that these hallmark properties of one or the other regime usually occur together.

Which orbits are "long?"

Note that all orbits in an n-dimensional Boolean network length at most 2^n. In the ordered regime, the length of orbits scales like a low-degree polynomial in n, in the chaotic regime, they scale exponentially in n.

For $1 < c < 2$, we call an n-dimensional Boolean network c-*chaotic* if it has at least one orbit of length $\geq c^n$.

We are interested in the question which conditions on the network preclude c-chaotic dynamics for c sufficiently close to 2.

p-fluid networks

In the ordered dynamics, along the attractors reached from most initial states, a large proportion of the variables will never change their values; such variables are called *eventually frozen.*

Let us call a Boolean network p-*fluid* if for a randomly chosen initial state with probability at least p a proportion of at most $1 - p$ of the variables are frozen.

We are interested in the question which conditions on the network preclude p-fluid dynamics for p sufficiently close to 1.

Homeostatic (in)stability

In the ordered regime, most single-bit flips in most initial conditions will leave the trajectory in the same basin of attraction. This property is called *high homeostatic stability.*

Let us call a Boolean system p-*unstable* if a random bit flip in a randomly chosen initial state with probability at least p moves the trajectory into the basin of attraction of a different attractor.

We are interested in the question which conditions on the network preclude p-unstable dynamics for p sufficiently close to 1.

Extremely chaotic networks

Suppose $c \approx 2$ and $p \approx 1$. A Boolean network that is simultaneously c-chaotic, p-fluid, and p-unstable should be considered *extremely chaotic.*

Which conditions on the network preclude extremely chaotic dynamics?

Regulatory functions

Recall that the dynamics of a Boolean network is defined by

$$s(t+1) = g(s(t)).$$

The function $g = [g_1, \ldots, g_n]$ is a Boolean vector function.

We call its i-th component g_i the *i-th regulatory function.* Ususally, g_i will depend on only some of the coordinates of $s = [s_1, \ldots, s_n]$; we call these the *inputs of variable i*.

If g_i depends on a coordinate s_j, then we say that i *is an output of variable j.*

Candidate conditions

Empirical studies of random Boolean networks have shown that the dynamics tends to be in the ordered regime if:

- All regulatory function have only a *small number of inputs*, or
- all regulatory functions are *nested canalyzing*, or
- all regulatory functions are *strongly biased*, or
- there are *few negative feedback loops.*

Do these conditions **provably** preclude extremely chaotic dynamics?

(b, r)-networks

A (b, r)-Boolean network is one in which each variable has at most r inputs and at most b outputs.

If $r = 2$, we call the system *quadratic;* a $(2, 2)$-system is called *bi-quadratic.* A regulatory function that depends on only one variable is called *monic;* a non-monic quadratic regulatory function is called *strictly quadratic.*

Nested canalyzing functions

The formal definition of nested canalyzing functions is a bit complicated, but examples of such functions include:

- $f(s) = s_i$,
- $f(s) = \neg s_i$,
- $f(s) = s_i \wedge s_j$,
- $f(s) = s_i \vee s_j$.

Biased Boolean functions

The *bias* Λ of a Boolean function is the proportion of input vectors on which the function takes the value 1.

For example, the bias of each monic function is 0.5. The bias of the strictly quadratic function $s_i \vee s_j$ is 0.75; the largest possible bias for any non-constant quadratic Boolean function.

We call a Boolean network ϵ-*biased* if $|\Lambda - 0.5| \geq \epsilon$ for each non-monic regulatory function.

Cooperative Boolean networks

Define the *cooperative (partial) order* on Π by $r \leq s$ if $r_i \leq s_i$ for $i = 1, \ldots, n$.

A Boolean network is *cooperative* if

$$r \leq s \to g(r) \leq g(s).$$

Note that negative feedback loops are totally absent in cooperative Boolean systems.

Cooperative quadratic Boolean networks

The only regulatory functions allowed in cooperative quadratic Boolean networks are $g_i = s_j$, $g_i = s_j \vee s_k$, $g_i = s_j \wedge s_k$.

Thus cooperative quadratic Boolean networks are 0.25-biased, have only nested canalyzing regulatory functions with few inputs and admit no negative feedback. Hence they satisfy all **empirical** conditions mentioned above.

Can one **prove** that such networks cannot have extremely chaotic dynamics?

Extreme chaos is still possible

Theorem 1. *Let c, p be constants with $1 < c < 2$ and $0 < p < 1$. Then for all sufficiently large n there exist n-dimensional Boolean networks that are simultaneously:*

(i) cooperative and bi-quadratic,

(ii) c-chaotic,

(iii) p-fluid,

(iv) p-unstable.

Idea of the proof

Fix $c < 2$. The idea is to set aside a small subset M of the variables to code a Turing machine that writes successive binary codes of integers $0, 1, \ldots, 2^{n-|M|}$ on a circular tape that is coded by the remaining variables. If M is of fixed size, this will give orbits of length $> c^n$ for sufficiently large n.

Unfortunately, this cannot be done in such a way that the resulting system is cooperative.

Overcoming the problem

Let L, ℓ be such that $\binom{L}{L/2} > 2^{\ell} > c^{L}$. Use L circular synchronously advancing tapes to code integers $0, 1, \ldots, 2^{n\ell/L}-1$ in such a way that the Turing machine M for computing the function $\oplus 1$ does not require negation. The size of M depends only on L.

This works. For sufficiently large n the construction yields a c-chaotic system, and after some further tweaking one that is also p-fluid and p-unstable.

Are there other examples?

The Turing machine metaphor comes readily to mind if a (former) logician wants to construct a counterexample. But Nature may have different ways of cobbling together gene regulatory networks. Are there examples of extremely chaotic, bi-quadratic, cooperative Boolean networks that look *radically different* from the ones we constructed?

Turing systems

In the construction described above, the vast majority of variables belong to the “tapes.” They have monic regulatory functions and simply copy previous values of other variables on the “tapes.”

Let us call an n-dimensional Boolean system an (M, n)-*Turing system* if at least $n - M$ of the regulatory functions are monic.

For $\alpha > 0$ and n sufficiently large, the example constructed in the proof of Theorem 1 is an $(\alpha n, n)$-Turing system.

Turing systems are the only examples

Theorem 2. *Let* $\epsilon, \alpha > 0$ *and let* b, r *be positive integers. Then there exists a positive constant* $c(\epsilon, \alpha, b, r) < 2$ *such that for every* $c > c(\epsilon, \alpha, b, r)$ *and sufficiently large* n*, every* c*-chaotic,* n*-dimensional* ϵ*-biased* (b, r)*-Boolean system is an* $(\alpha n, n)$*-Turing system.*

Note that we do not need to assume that the system is cooperative.

We will let $c(\epsilon, \alpha, b, r)$ denote the smallest number for which the conclusion of Theorem 2 holds.

Interpreting $c(0.25, \alpha, 2, 2)$

If an n-dimesional bi-quadratic cooperative Boolean network has an orbit of size $\geq c^n$ for some $c > c(0.25, 0.1, 2, 2)$, then at most 10% of all variables in the system have strictly quadratic regulatory functions.

Similarly, if the system has an orbit of size $\geq c^n$ for some $c > c(0.25, 1, 2, 2)$, then at least some variable in the system has a monic regulatory function.

Estimates for $c(0.25, \alpha, 2, 2)$

We can show that

$$c(0.25, \alpha, 2, 2) \leq 10^{(2-\alpha)/4} \qquad (1)$$

for all $\alpha \in [0, 1]$.

If $\alpha = 1$, this gives $c(0.25, \alpha, 2, 2) \leq 10^{1/4}$, and we can also show that equality holds in this case.

We can also show that

$$c(0.25, \alpha, 2, 2) \leq 2 - 0.0041\alpha, \qquad (2)$$

which is a sharper bound than (1) for $\alpha < 0.7987$.

Is "bi"-quadratic necessary?

Can we prove an analogue of Theorem 2 if no bound on the number of outputs per variable is assumed? No!

Theorem 3. *Let c, p be constants with $1 < c < 2$ and $0 < p < 1$. Then for all sufficiently large n there exist n-dimensional Boolean networks that are simultaneously:*

(i) cooperative and have only strictly quadratic regulatory functions,

(ii) c-chaotic,

(iii) p-fluid,

(iv) p-unstable.

Can we prove an analogue of Theorem 2 for p-instability?

No!

Theorem 4. *Let n be a positive integer. Then there exists a* 1*-unstable, strictly bi-quadratic cooperative Boolean system of dimension* $2n$.

How about extreme chaos in its full glory?

Theorem 5. *Let* c *be a constant such that* $2\sqrt{0.75} < c < 2$ *and let* $p > 0.75 + \frac{\ln(0.5c)}{2 \ln 0.75}$. *Then no strictly quadratic cooperative Boolean system can simultaneously be* c*-chaotic and* p*-unstable.*

For example, no sufficiently high-dimensional strictly quadratic cooperative Boolean network can simultaneously be be 0.9-unstable and 1.85-chaotic.

Open problem

Determine the exact values of $c(\epsilon, \alpha, b, r)$ when $[\epsilon, \alpha, b, r] \neq [0.25, 1, 2, 2]$.