Attractor stucture of Boolean networks of small canalizing depth

Gleb Pogudin,
MAX team, LIX, CNRS, École Polytechnique, Institut Polytechnique de Paris,

joint work with R. Laubenbacher, E. Paul, and W. Qin

INSTITUT
POLYTECHNIQUE
DE PARIS

Plan of the talk

Main question: difference in dynamics between
"a random Boolean network" and "random canalizing Boolean network"

Plan

- Intro and problem statement
- Simulation results
- Theoretic results
- Remarks on the computation

Boolean networks and their structure

Definition

$\mathbf{F}_2$ is a field with two elements 0 (false / off) and 1 (true / on) and arithmetic operations:

$$x + y := x \text{ XOR } y, \quad xy := x \text{ AND } y.$$

Boolean networks and their structure

Definition

$\mathbf{F}_2$ is a field with two elements 0 (false / off) and 1 (true / on) and arithmetic operations:

$$x + y := x \text{ XOR } y, \quad xy := x \text{ AND } y.$$

An n-dimensional Boolean network defined by n polynomials $f_1, \ldots, f_n \in \mathbb{F}_2[x_1, \ldots, x_n]$ is a discrete-time dynamical system such that

- states are points in $\mathbb{F}_2^n$
- and, the the state at time t is $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,n}) \in \mathbb{F}_2^n$, then

$$\mathbf{a}_{t+1} = (f_1(\mathbf{a}_t), \ldots, f_n(\mathbf{a}_t)).$$

Boolean networks and their structure

Definition

$\mathbf{F}_2$ is a field with two elements 0 (false / off) and 1 (true / on) and arithmetic operations:

$$x + y := x \text{ XOR } y, \quad xy := x \text{ AND } y.$$

An n-dimensional Boolean network defined by n polynomials $f_1, \ldots, f_n \in \mathbb{F}_2[x_1, \ldots, x_n]$ is a discrete-time dynamical system such that

- states are points in $\mathbb{F}_2^n$
- and, the the state at time t is $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,n}) \in \mathbb{F}_2^n$, then

$$\mathbf{a}_{t+1} = (f_1(\mathbf{a}_t), \ldots, f_n(\mathbf{a}_t)).$$

Example

Let $f_1(x_1, x_2) = x_1(x_2 + 1)$ and $f_2(x_1, x_2) = (x_1 + 1)x_2$:

$$(0, 1) \to (1, 0) \to (0, 1) \to \ldots$$

Boolean networks and their structure

Definition

An n-dimensional Boolean network defined by n polynomials $f_1, \ldots, f_n \in \mathbb{F}_2[x_1, \ldots, x_n]$ is a discrete-time dynamical system such that

- states are points in $\mathbb{F}_2^n$
- and, the the state at time t is $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,n}) \in \mathbb{F}_2^n$, then

$$\mathbf{a}_{t+1} = (f_1(\mathbf{a}_t), \ldots, f_n(\mathbf{a}_t)).$$

Why: regulatory networks.

Boolean networks and their structure

Definition

An n-dimensional Boolean network defined by n polynomials $f_1, \ldots, f_n \in \mathbb{F}_2[x_1, \ldots, x_n]$ is a discrete-time dynamical system such that

- states are points in $\mathbb{F}_2^n$
- and, the the state at time t is $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,n}) \in \mathbb{F}_2^n$, then

$$\mathbf{a}_{t+1} = (f_1(\mathbf{a}_t), \ldots, f_n(\mathbf{a}_t)).$$

Why: regulatory networks.

Why polynomials?

$\mathbb{F}^n$ is a finite set $\implies$ any function is polynomial

Boolean networks and their structure

Definition

An n-dimensional Boolean network defined by n polynomials $f_1, \ldots, f_n \in \mathbb{F}_2[x_1, \ldots, x_n]$ is a discrete-time dynamical system such that

- states are points in $\mathbb{F}_2^n$
- and, the the state at time t is $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,n}) \in \mathbb{F}_2^n$, then

$$\mathbf{a}_{t+1} = (f_1(\mathbf{a}_t), \ldots, f_n(\mathbf{a}_t)).$$

Why: regulatory networks.

Why polynomials?

$\mathbb{F}^n$ is a finite set $\implies$ any function is polynomial

Pros: we can define any dynamics!

Boolean networks and their structure

Definition

An n-dimensional Boolean network defined by n polynomials $f_1, \ldots, f_n \in \mathbb{F}_2[x_1, \ldots, x_n]$ is a discrete-time dynamical system such that

- states are points in $\mathbb{F}_2^n$
- and, the the state at time t is $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,n}) \in \mathbb{F}_2^n$, then

$$\mathbf{a}_{t+1} = (f_1(\mathbf{a}_t), \ldots, f_n(\mathbf{a}_t)).$$

Why: regulatory networks.

Why polynomials?

$\mathbb{F}^n$ is a finite set $\implies$ any function is polynomial

Pros: we can define any dynamics!

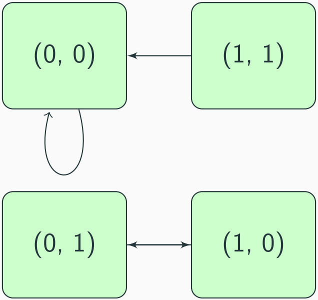
Cons: we can define any dynamics...

Dynamics: attractors and basins

n-dimensional Boolean network $\implies$ directed graph on 2^n vertices

Network:

$$f_1 = x_1(x_2 + 1),$$
$$f_2 = (x_1 + 1)x_2.$$

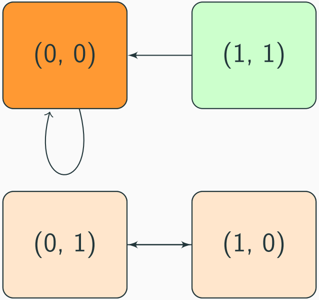


Dynamics: attractors and basins

n-dimensional Boolean network $\implies$ directed graph on 2^n vertices

Network:

$$f_1 = x_1(x_2 + 1),$$
$$f_2 = (x_1 + 1)x_2.$$



Notions

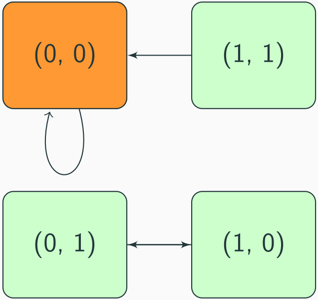
- Attractors are cycles in the graph.

Dynamics: attractors and basins

n-dimensional Boolean network $\Longrightarrow$ directed graph on 2^n vertices

Network:

$$f_1 = x_1(x_2 + 1),$$
$$f_2 = (x_1 + 1)x_2.$$



Notions

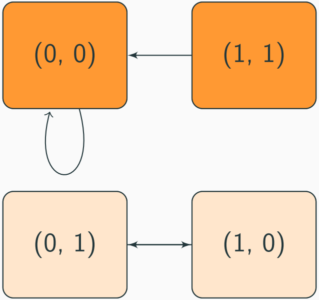
- Attractors are cycles in the graph.
- Steady states are loops in the graph.

Dynamics: attractors and basins

n-dimensional Boolean network $\Longrightarrow$ directed graph on 2^n vertices

Network:

$$f_1 = x_1(x_2 + 1),$$
$$f_2 = (x_1 + 1)x_2.$$



Notions

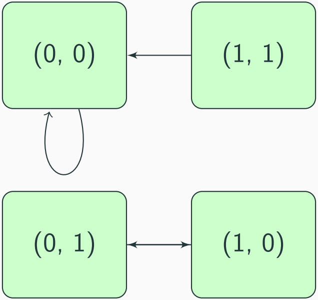
- Attractors are cycles in the graph.
- Steady states are loops in the graph.
- Basins are connected components.
  Each basin corresponds to an attractor and each attractor has its basin.

Dynamics: attractors and basins

n-dimensional Boolean network $\implies$ directed graph on 2^n vertices

Network:

$$f_1 = x_1(x_2 + 1),$$
$$f_2 = (x_1 + 1)x_2.$$



Notions

- Attractors are cycles in the graph.
- Steady states are loops in the graph.
- Basins are connected components.
  Each basin corresponds to an attractor and each attractor has its basin.

Intuition
Attractors $\leftrightarrow$ Cell types
Length $\leftrightarrow$ number of experssion patterns

Canalization

One variable to rule them all

A nonconstant function $f(x_1, \dots x_n)$ is canalizing with respect to a variable x_i if there exists a canalizing value $a \in \{0, 1\}$ such that

$$f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \equiv \text{const}.$$

Canalization

One variable to rule them all

A nonconstant function $f(x_1, \dots x_n)$ is canalizing with respect to a variable x_i if there exists a canalizing value $a \in \{0, 1\}$ such that

$$f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \equiv \text{const.}$$

Example

$f(x_1, x_2) = x_1(x_2 + 1)$ is canalizing w.r.t

x_1 with value 0

x_2 with value 1

Canalization

One variable to rule them all

A nonconstant function $f(x_1, \ldots x_n)$ is canalizing with respect to a variable x_i if there exists a canalizing value $a \in \{0, 1\}$ such that

$$f(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n) \equiv \text{const.}$$

Example

$f(x_1, x_2) = x_1(x_2 + 1)$ is canalizing w.r.t

x_1 with value 0

x_2 with value 1

Why?

- Often appear in the published models.
- Make sense.

Canalization

One variable to rule them all

A nonconstant function $f(x_1, \dots x_n)$ is canalizing with respect to a variable x_i if there exists a canalizing value $a \in \{0, 1\}$ such that

$$f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \equiv \text{const.}$$

Canalizing depth (Layne, Dimitrova, Macaulay, 2012)

We define canalizing depth recursively

1. noncanalizing function $\implies$ depth 0
2. if f is canalizing w.r.t. x_1 with value a, then

$$\text{depth}(f) = \text{depth}(f|_{x_1=1-a}) + 1.$$

Main question

Question: How does the canalization (and its depth) affect the attractor structure?

Main question

Question: How does the canalization (and its depth) affect the attractor structure?

Related results include: no chaos (Karlsson, Hörnquist), stability (Kauffman, Peterson, Samuelsson, Troein and Layne, Dimitrova, Macauley)

Experimental results

Setup

For each $n = 4, \ldots, 20$ and each $k = 0, \ldots, n$ we generate random Boolean networks

- of dimension n;
- with all the functions having canalizing depth k.

Experimental results

Setup

For each $n = 4, \dots, 20$ and each $k = 0, \dots, n$ we generate random Boolean networks

- of dimension n;
- with all the functions having canalizing depth k.

Computed values

For each Boolean network we compute

N: the number of attractors

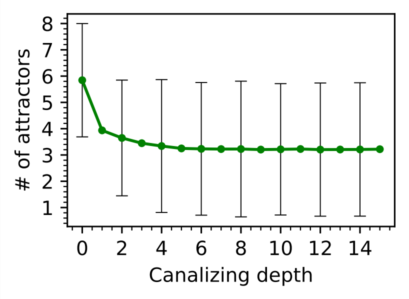
S: the total number of states in the attractors

AS: the average attractor size $= \frac{S}{N}$

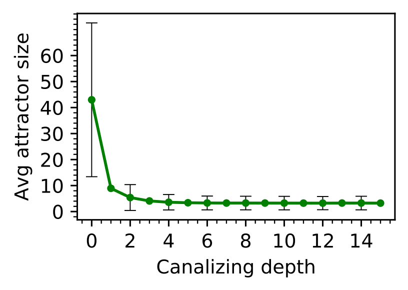
Experimental results

For $n = 15$ based in 50k samples for each k.

N = # attractors



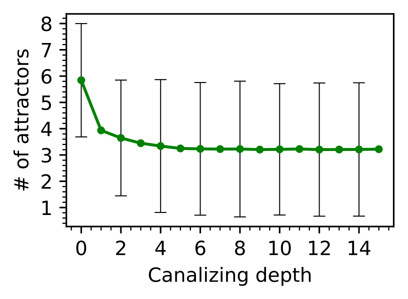
AS = # avg. attractor size



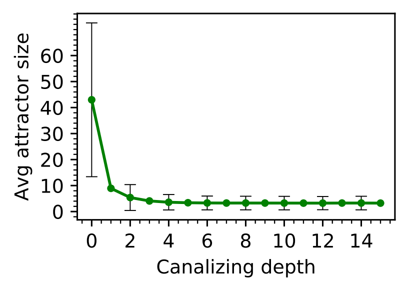
Experimental results

For $n = 15$ based in 50k samples for each k.

N = # attractors



AS = # avg. attractor size



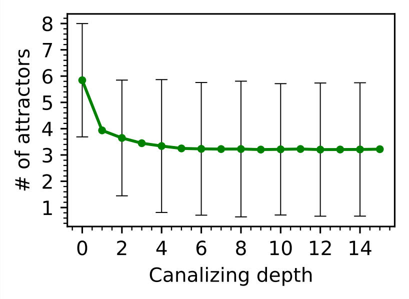
Analysis

- Depth increases $\implies$ number and sizes decrease

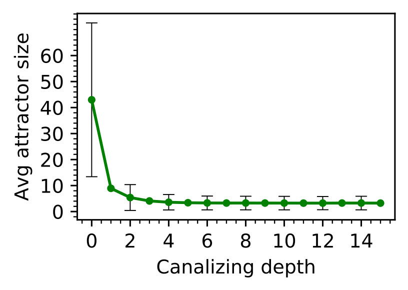
Experimental results

For $n = 15$ based in 50k samples for each k.

N $=$ # attractors



AS $=$ # avg. attractor size



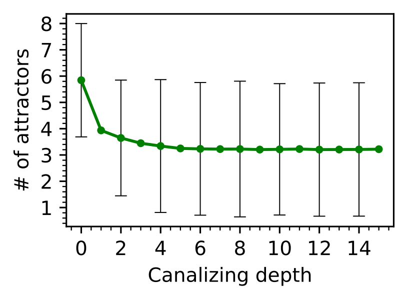
Analysis

- Depth increases $\implies$ number and sizes decrease
- Sizes decrease more substantially

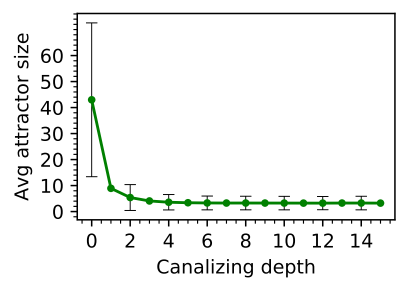
Experimental results

For $n = 15$ based in 50k samples for each k.

N $=$ # attractors



AS $=$ # avg. attractor size



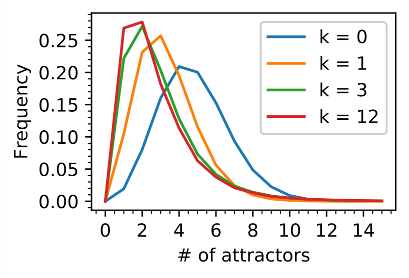
Analysis

- Depth increases $\implies$ number and sizes decrease
- Sizes decrease more substantially
- Most of the change happens between $k = 0$ and $k = 3$.

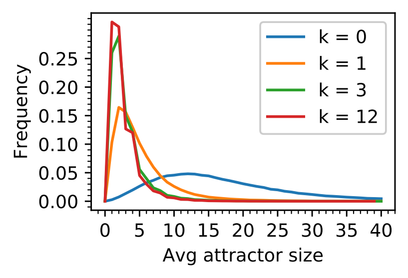
Experimental results

For $n = 12$ based in 300k samples for each k.

N = # attractors



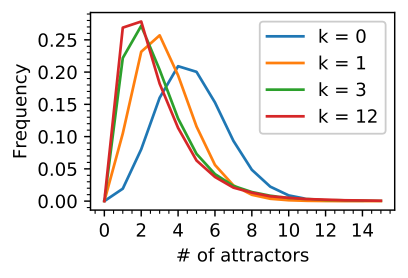
AS = # avg. attractor size



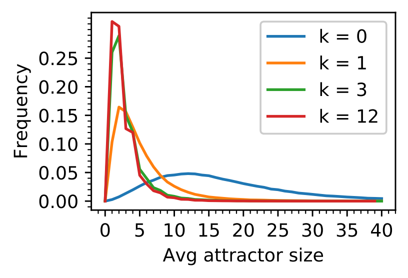
Experimental results

For $n = 12$ based in 300k samples for each k.

N $= \#$ attractors



AS $= \#$ avg. attractor size



Analysis

- Confirms earlier conclusions
- Avg. size differs dramatically for $k = 0$ and $k = 1$

Theoretical results

Quantity of interest

$A_{\ell,n}$:= expected # of attractors of length ℓ in a random network of dimension n and depth one.

$$A_\ell := \lim_{n\to\infty} A_\ell$$

Theoretical results

Quantity of interest

$A_{\ell,n} :=$ expected # of attractors of length ℓ in a random network of dimension n and depth one.

$$A_\ell := \lim_{n\to\infty} A_\ell$$

What if depth $= 0$?
The corresponding limit will be $\frac{1}{\ell}$

Theoretical results

Quantity of interest

$A_{\ell,n} :=$ expected # of attractors of length ℓ in a random network of dimension n and depth one.

$$A_\ell := \lim_{n\to\infty} A_\ell$$

What if depth $= 0$?
The corresponding limit will be $\frac{1}{\ell}$

Theorem

We give a formula for A_ℓ, in particular

- $A_\ell > \frac{1}{\ell}$ for $\ell > 0$;
- $A_1 = 1,\ A_2 = 0.66\ldots,\ A_3 = 0.3386\ldots,\ A_4 = 0.2856\ldots$

Summary

- canalization $\implies$ fewer attractors and they are much smaller
  Biologically relevant!

Summary

- canalization $\implies$ fewer attractors and they are much smaller
  Biologically relevant!
- no big difference between small (but > 0) and large canalizing depths
  Applicable to a wide range of models

Remarks on the computation

Sampling:

- Relies on the structure theorem by He and Macaulay
- Complexity $\mathcal{O}(n2^n)$ almost optimal
  more precisely $\mathcal{O}(k^3 + (n-k)2^{n-k} + k2^n)$

Finding attractors: DFS

Open problems

Taking into account:

- sparsity
- uneven canalization profiles
  (i.e. only some of the functions are canalizing)
- collective canalization

Support

This work was partially supported by the National Science Foundation.