Network Structure and Activity in Boolean Networks

Henning S. Mortveit

with

Abhijin Adiga, Hilton Galyean, Chris J. Kuhlman,
Michael Levet, and Sichao Wu

Department of Mathematics &
Network Dynamics and Simulation Science Laboratory, VBI
Virginia Tech

Automata 2015, Turku Finland, June 8–10, 2015

Boolean Networks (BN)

- A Boolean Network (BN) is a map of the form

$$F = (f_1, \ldots, f_n)\colon \{0,1\}^n \longrightarrow \{0,1\}^n .$$

- There are several types of stability concepts for dynamical systems. Examples:
  - Stability with respect to initial conditions
  - Attractor stability
  - Structural stability

Boolean Networks (BN)

- A Boolean Network (BN) is a map of the form

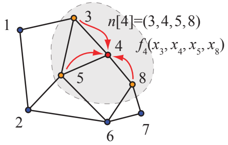
$$F = (f_1, \ldots, f_n)\colon \{0,1\}^n \longrightarrow \{0,1\}^n .$$

- There are several types of stability concepts for dynamical systems. Examples:
  - Stability with respect to initial conditions
  - Attractor stability
  - Structural stability
- Here: Stability with respect to initial condictions. Specifically:

What is the probability that $F(x)$ and $F(x + e_i)$ are different?

Background, Terminology and Definitions – I

- Associated to F we have the *dependency graph of variables* denoted by X_F. It has *vertex set* $\{1, 2, \ldots, n\}$ and *edges* all (i, j) for which f_j depend non-trivially on x_j.
- We let X denote the undirected version of X_F where loops are omitted.
- More definitions:
  - Ordered 1-neighborhood: $n[i]$
  - System state: $x = (x_1, \ldots, x_n)$
  - Restricted vertex state: $x[i]$



Background, Terminology and Definitions – II

- Define the function $\alpha_{F,i}\colon K^n \longrightarrow \{0,1\}$ by

$$\alpha_{F,i}(x) = \mathbb{I}[F(x+e_i) \neq F(x)]$$

where $\mathbb{I}$ is the indicator function and e_i is the i^{th} unit vector.

Definition

The *activity of F with respect to vertex i* is the expectation value of $\alpha_{F,i}$ using the uniform measure on K^n:

$$\bar{\alpha}_{F,i} = \mathbb{E}[\alpha_{F,i}]\ .$$

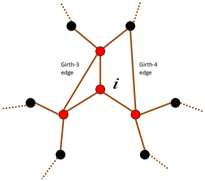
The *activity of F* is the vector

$$\bar{\alpha}_F = (\bar{\alpha}_{F,1}, \bar{\alpha}_{F,2}, \ldots, \bar{\alpha}_{F,n})\ .$$

Comments and Definitions

- Interpretation: For a randomly chosen state $x \in K^n$, the value $\bar{\alpha}_{F,i}$ may be regarded as the probability that perturbing x_i will cause $F(x+e_i) \neq F(x)$ to hold. This activity notion may naturally be regarded as a measure of sensitivity with respect to initial conditions.

  The subgraph $X(i;2)$ of X is induced by vertex i and its distance ≤ 2 neighbors.



- Vertices belonging to the closed 1-neighborhood $n[i]$ of i are marked red. *Type-3 edges* (relative to i) connect neighbors of i, while *type-4* edges connect neighbors of $j \in n'[i]$ through a common neighbor different from i.

- Observation: $\bar{\alpha}_{F_i}(x)$ depends on the vertex states $X(i;2)$ and the vertex functions f_j with $j \in n[i]$ (the set of vertices in the 1-neighborhood of i).

Related Work

- Shmulevich and Kauffman (2003) introduced a notion of activity for regular, random Boolean networks with a common vertex function f (see [3]). Here:

$$\alpha_{f,i} := \mathbb{E}[\frac{\partial f}{\partial x_i}]$$

- May be a reasonable measure for activity for the class of dynamical systems that they studied. However: does not incorporate the impact of network structure.

- Layne et. al. (2012) used the above definition to compute activity for canalyzing functions [2].

- The activity notion introduced here incorporates network structure.

Related Work

- Shmulevich and Kauffman (2003) introduced a notion of activity for regular, random Boolean networks with a common vertex function f (see [3]). Here:

$$\alpha_{f,i} := \mathbb{E}[\frac{\partial f}{\partial x_i}]$$

- May be a reasonable measure for activity for the class of dynamical systems that they studied. However: does not incorporate the impact of network structure.
- Layne et. al. (2012) used the above definition to compute activity for canalyzing functions [2].
- The activity notion introduced here incorporates network structure.
- Why study activity? What are potential uses?

Preliminary Results

- We set $K = \{0,1\}$, write N_i for the size of $X(i;2)$, and $K(i) = K^{N_i}$ for the projection of K^n onto the set of vertex states associated to $X(i;2)$.
- For $j \in n[i]$, define the sets $A_j(i) \subset K(i)$ and $A_j^0, A_j^1 \subset A_j(i)$ by

$$A_j(i) = \{x \in K(i) \mid F(x+e_i)_j \neq F(x)_j\} \quad \text{and} \quad A_j^m(i) = \{x \in A_j(i) \mid x_i = m\}\ , m = 0, 1\ .$$

Preliminary Results

- We set $K = \{0,1\}$, write N_i for the size of $X(i;2)$, and $K(i) = K^{N_i}$ for the projection of K^n onto the set of vertex states associated to $X(i;2)$.
- For $j \in n[i]$, define the sets $A_j(i) \subset K(i)$ and $A_j^0, A_j^1 \subset A_j(i)$ by

$$A_j(i) = \{x \in K(i) \mid F(x+e_i)_j \neq F(x)_j\} \quad \text{and} \quad A_j^m(i) = \{x \in A_j(i) \mid x_i = m\}\ , m = 0, 1\ .$$

Proposition

For any Boolean network we have

$$\mathbb{P}\Big(\bigcup_{j\in n[i]} A_j \mid x_i = 0\Big) = \mathbb{P}\Big(\bigcup_{j\in n[i]} A_j \mid x_i = 1\Big)\ .$$

Proof.

This follows by observing that there is a bijection $\phi_j \colon A_j^0 \longrightarrow A_j^1$ since $x \in A_j^0$ if and only if $x + e_i \in A_j^1$ for all $j \in n[i]$.

Proposition

Let X be a graph and F a Boolean network map over X. The activity of F with respect to vertex i is

$$\bar{\alpha}_{F,i} = \mathbb{P}\Big(\bigcup_{j \in n[i]} A_j \mid x_i = 0 \Big) = \mathbb{P}\Big(\bigcup_{j \in n[i]} \bar{A}_j \Big) .$$

(Here $\bar{A}_j := A_j^0$)

Proposition

Let X be a graph and F a Boolean network map over X. The activity of F with respect to vertex i is

$$\bar{\alpha}_{F,i} = \mathbb{P}\Big(\bigcup_{j \in n[i]} A_j \mid x_i = 0 \Big) = \mathbb{P}\Big(\bigcup_{j \in n[i]} \bar{A}_j \Big) \ .$$

Proof.

$$\bar{\alpha}_{F,i} = \mathbb{E}[\alpha_{F,i}] = \sum_{x \in K^n} \alpha_{F,i}(x)\mathbb{P}(x) = \sum_{x \in K(i)} \alpha_{F,i}(x)\mathbb{P}(x) = \mathbb{P}\Big[\bigcup_{j \in n[i]} A_j \Big]$$

$$= \sum_{m=0,1} \mathbb{P}\Big[\bigcup_{j \in n[i]} A_j \mid x_i = m \Big] \mathbb{P}(x_i = m) = \mathbb{P}\Big[\bigcup_{j \in n[i]} A_j \mid x_i = 0 \Big] = \mathbb{P}\Big[\bigcup_{j \in n[i]} \bar{A}_j \Big] \ .$$

(Here $\bar{A}_j := A_j^0$)

Examples

Recall that the standard Boolean *threshold function* $\tau_{k,n}\colon K^n \longrightarrow K$ is defined by

$$\tau_{k,n}(x_1,\ldots,x_n) = \begin{cases} 1, & \text{if } \sum_{j=1}^{n} x_j \geq k, \\ 0, & \text{otherwise,} \end{cases}$$

and that the logical nor-function $\mathrm{nor}_m\colon K^m \longrightarrow K$ is defined by

$$\mathrm{nor}_m(x_1,\ldots,x_m) = (1+x_1)\cdot\cdots\cdot(1+x_m)\ ,$$

with arithmetic operations modulo 2.

Examples

Recall that the standard Boolean *threshold function* $\tau_{k,n}\colon K^n \longrightarrow K$ is defined by

$$\tau_{k,n}(x_1,\ldots,x_n) = \begin{cases} 1, & \text{if } \sum_{j=1}^{n} x_j \geq k, \\ 0, & \text{otherwise,} \end{cases}$$

and that the logical nor-function $\mathrm{nor}_m\colon K^m \longrightarrow K$ is defined by

$$\mathrm{nor}_m(x_1,\ldots,x_m) = (1+x_1)\cdot\cdots\cdot(1+x_m)\ ,$$

with arithmetic operations modulo 2.

Proposition

If F is the BN map induced by the nor*-function over K_n, then*

$$\bar{\alpha}_{F,i} = \frac{1}{2^{n-1}},$$

and if F is induced by the k-threshold function over K_n, then

$$\bar{\alpha}_{F,i} = \binom{n-1}{k-1}\Big/ 2^{n-1}\ .$$

Proof.

For either choice of function, we have $\bar{A}_j = \bar{A}_k$ for all $j, k \in n[i]$.

For threshold functions, and for a state x with $x_i = 0$ to satisfy $\tau_k(x) \neq \tau_k(x + e_i)$, it is necessary and sufficient that x belong to Hamming class $k - 1$.

Since $x_i = 0$, it follows from Proposition 2 that $\mathbb{P}(\bar{A}_j(i)) = \frac{1}{2^{n-1}}|A_j| = \binom{n-1}{k-1}/2^{n-1}$ as stated.

In the case of nor-functions we have $F\left(x + e_i\right)_i \neq F\left(x\right)_i$ precisely when $x_j = 0$ for all $j \in n(i)$ leading to $\mathbb{P}\left(\bar{A}_i\right) = \frac{1}{2^{n-1}}$.

Elementary Cellular Automata

Proposition (Threshold ECA)

The activity of k-threshold ECA for $n \geq 5$ is

$$\bar{\alpha}_{F,i} = \begin{cases} 0, & \text{if } k = 0 \text{ or } k > 3\,, \\ 1/2, & \text{if } k = 1 \text{ or } k = 3\,, \\ 7/8, & \text{if } k = 2\,. \end{cases}$$

For $n = 4$ and $k = 2$, the activity is $3/4$.

Proof (case $k = 1$).

We have

$$\bar{A}_{i-1} = \{(0,0,0,x_{i+1},x_{i+2})\}, \quad \bar{A}_i = \{(x_{i-2},0,0,0,x_{i+2})\}, \text{ and } \bar{A}_{i+1} = \{(x_{i-1},x_{i-1},0,0,0)\} .$$

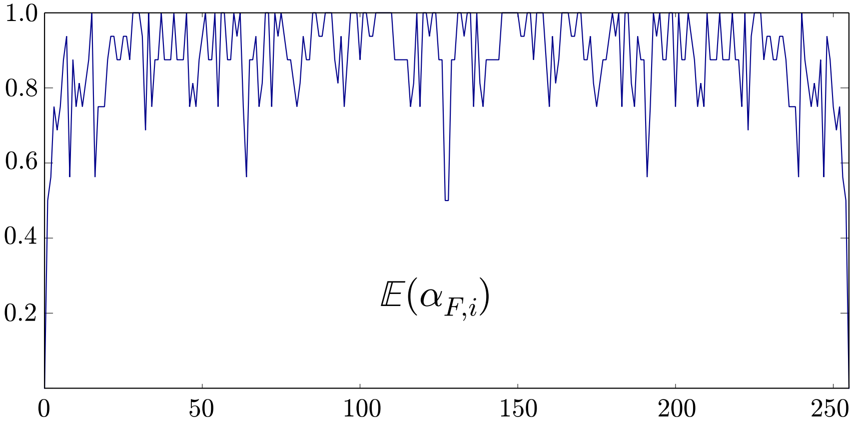
By the inclusion-exclusion principle:

$$\begin{aligned} |\bar{A}_{i-1} \cup \bar{A}_i \cup \bar{A}_{i+1}| &= |\bar{A}_{i-1}| + |\bar{A}_i| + |\bar{A}_{i+1}| \\ &\quad - |\bar{A}_{i-1} \cap \bar{A}_i| - |\bar{A}_{i-1} \cap \bar{A}_{i+1}| - |\bar{A}_i \cap \bar{A}_{i+1}| + |\bar{A}_{i-1} \cap \bar{A}_i \cap \bar{A}_{i+1}| \\ &= 3 \times 4 - 2 - 1 - 2 + 1 = 8 \end{aligned}$$

This yields $\bar{\alpha}_{F,i} = 8/2^4 = 1/2$.

We remark that if we instead use the nor-function, then $\bar{\alpha}_{F,i} = 1/2$ when $n \geq 5$.

Activity of Elementary Cellular Automata



More Elaborate Cases

- The evaluation of activity is more involved with dependencies/overlap among the A_j's (except for at vertex i – of course).
- Will cover two cases here:
  - d-regular trees with threshold and nor-functions
  - Square lattices with nor-functions
- Additional cases covered in the conference paper.

d-regular trees

Proposition

Let X be a d-regular graph of girth ≥ 5 and F the GDS map over X induced by the nor*-function. Then the activity of F with respect to i is given by*

$$\bar{\alpha}_{F,i} = 1 - \left(1 - \frac{1}{2^d}\right)^d + \left(\frac{1}{2} - \frac{1}{2^d}\right)^d \ . \quad (1)$$

Note: let B denote the union of A_j's with $j \neq i$. Then:

$$\mathbb{P}\left[\bigcup_{j \in n[i]} A_j\right] = \mathbb{P}(A_i \cup B) = \mathbb{P}(B) + \mathbb{P}(A_i \cap B^c) = 1 - \mathbb{P}(B^c) + \mathbb{P}(A_i \cap B^c) \ ,$$

where B^c denotes the complement of B.

Proof.

Conditioning on $x_i = 0$ and using independence we have

$$\mathbb{P}\left(\bar{B}^c\right) = \mathbb{P}\left(\bigcap_{j \in n(i)} \bar{A}_j^c\right) = \prod_{j \in n(i)} \mathbb{P}\left(\bar{A}_j\right) \ .$$

For a state x with $x_i = 0$ to be in $\bar{A}_j$ all remaining d states of $\bar{A}_j$ must be zero, leading to $\mathbb{P}\left(\bar{A}_j^c\right) = 1 - \frac{1}{2^d}$ and

$$\mathbb{P}\left(\bar{B}^c\right) = \left(1 - \frac{1}{2^d}\right)^d \ .$$

Proof.

Conditioning on $x_i = 0$ and using independence we have

$$\mathbb{P}\left(\bar{B}^c\right) = \mathbb{P}\left(\bigcap_{j\in n(i)} \bar{A}_j^c\right) = \prod_{j\in n(i)} \mathbb{P}\left(\bar{A}_j\right) \ .$$

For a state x with $x_i = 0$ to be in $\bar{A}_j$ all remaining d states of $\bar{A}_j$ must be zero, leading to $\mathbb{P}\left(\bar{A}_j^c\right) = 1 - \frac{1}{2^d}$ and

$$\mathbb{P}\left(\bar{B}^c\right) = \left(1 - \frac{1}{2^d}\right)^d \ .$$

In order to calculate $\mathbb{P}\left(A_i \cap B^c\right)$ we note that $\mathbb{P}\left(A_i \cap B^c\right) = \mathbb{P}\left(A_i\right)\mathbb{P}\left(B^c|A_i\right)$ and again use independence to obtain

$$\mathbb{P}\left(\bar{B}^c|\bar{A}_i\right) = \prod_{j\in n(i)} \mathbb{P}\left(\bar{A}_j^c|\bar{A}_i\right) \ .$$

Note that $\mathbb{P}\left(\bar{A}_j|\bar{A}_i\right) = 1/2^{d-1}$ so that $\mathbb{P}\left(\bar{A}_j^c|\bar{A}_i\right) = 1 - \mathbb{P}\left(\bar{A}_j|\bar{A}_i\right) = 1 - 1/2^{d-1}$ which substituted back in gives

$$\mathbb{P}\left(\bar{B}^c|\bar{A}_i\right) = \left(1 - \frac{1}{2^{d-1}}\right)^d \ .$$

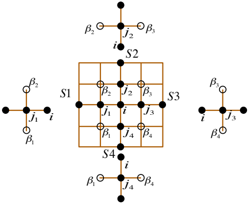
Noting that $\mathbb{P}(\bar{A}_i) = 1/2^d$ we obtain the third term in the proposition statement, finishing the proof

Proposition

If F is the GDS map with k-threshold vertex functions over a d-regular graph of girth ≥ 5, then the activity of F with respect to vertex i is

$$\bar{\alpha}_{F,i} = 1 - \left[\frac{2^d - \binom{d}{k-1}}{2^d}\right]^d + \frac{\binom{d}{k-1}}{2^d}\left[\frac{2^{d-1} - \binom{d-1}{k-2}}{2^{d-1}}\right]^{k-1}\left[\frac{2^{d-1} - \binom{d-1}{k-1}}{2^{d-1}}\right]^{d-(k-1)} .$$

Square lattices (type-4 edge-)



Proposition

Let X be the 2-dimensional lattice (infinite or periodic boundaries) as above where every vertex has degree 4, and let F be the GDS map over X induced by nor-functions. Then the activity of F for any vertex i is $\bar{\alpha}_{F,i} = 1040/2^{12}$.

Proof.

Idea: use inclusion-exclusion and careful book-keeping twice to evaluate $\mathbb{P}(B)$ and $\mathbb{P}(\bar{A}_i \cap \bar{B}^c)$.

Remark

May of course obtain this through brute force evaluation: only 2^{13} state involved for $X(i;2)$.

Questions, Acknowledgements and References

- Computational intractability: in the general case, determination of activity is intractable.
- Are there efficient and accurate methods for estimation?
- Activity only covers 1-step effects. Can something be done for long-term activity (e.g. with comparison of limit sets $\omega_F(x) \neq \omega_F(x + e_i)$?)
- Type-3 edges determines the [vertex] cluster coefficient. Can the cluster coefficient say something about activity?

Questions, Acknowledgements and References

- Computational intractability: in the general case, determination of activity is intractable.
- Are there efficient and accurate methods for estimation?
- Activity only covers 1-step effects. Can something be done for long-term activity (e.g. with comparison of limit sets $\omega_F(x) \neq \omega_F(x + e_i)$?)
- Type-3 edges determines the [vertex] cluster coefficient. Can the cluster coefficient say something about activity?

Acknowledgements. We thank our external collaborators and members of the Network Dynamics and Simulation Science Laboratory (NDSSL) for their suggestions and comments. This work has been partially supported by DTRA Grant HDTRA1-11-1-0016.

A. Adiga, H. Galyean, C. J. Kuhlman, M. Levet, H. S. Mortveit, and S. Wu, *Network structure and activity in boolean networks*.
Accepted in Proceedings of Automata 2015.

L. Layne, E. Dimitrova, and M. Matthew, *Nested canalyzing depth and network stability*, Bulletin of Mathematical Biology, 74 (2012).

I. Shmulevich and S. A. Kauffman, *Activities and sensitivities in Boolean network models*, Physical Review Letters, 93 (2004), pp. 048701:1–4.