Biología Matemática Discreta

David Murrugarra

Department of Mathematics,
University of Kentucky.

Grupo de Investigatión
Ciencias Matemáticas
Para las Ciencias de la Vida.
Lima, Perú, Mayo 17, 2018.

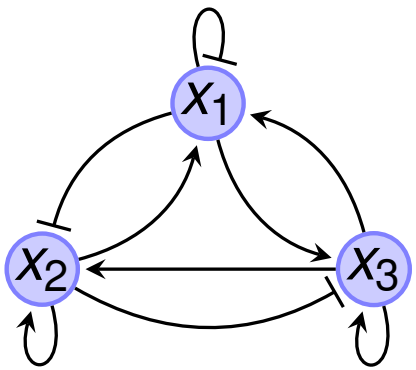
Discrete Dynamical Systems

$$\mathbf{F} = (f_1, \ldots, f_n) : X \to X$$

where

- $X = X_1 \times \cdots \times X_n$ is the Cartesian product of finite sets.
- $f_i : X \to X_i$ is the update function for x_i, for all $i = 1, \ldots, n$.
- Dynamics is generated by iteration,

$$x(t+1) = \mathbf{F}(x(t))$$



Network

Functions:

$$\begin{cases} x_1^{t+1} = f_1(x_1^t, x_2^t, x_3^t) \\ x_2^{t+1} = f_2(x_1^t, x_2^t, x_3^t) \\ x_3^{t+1} = f_3(x_1^t, x_2^t, x_3^t) \end{cases}$$

Gene Regulatory Networks: The *Lac* operon

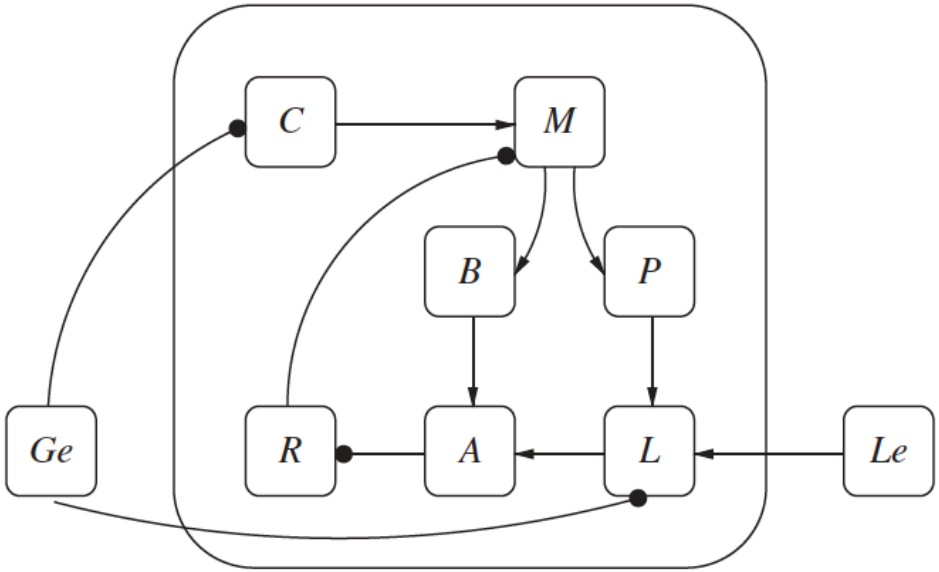
Genes

$x_1 = M$: *lac* mRNA,
$x_2 = P$: *lac permease*,
$x_3 = B$: *lac*β*-galactosidase*,
$x_4 = C$: CAP,
$x_5 = R$: repressor,
$x_6 = R_m$: medium repressor,
$x_7 = A$: allolactose,
$x_8 = A_m$: medium allolactose,
$x_9 = L$: lactose,
$x_{10} = L_m$: medium lactose,

Update functions

$f_1 = x_4 \wedge \overline{x_5}$,
$f_2 = x_1$,
$f_3 = x_1$,
$f_4 = \overline{G_e}$,
$f_5 = \overline{x_7} \wedge \overline{x_8}$,
$f_6 = (\overline{x_7} \wedge \overline{x_8}) \vee x_5$,
$f_7 = x_9 \wedge x_3$,
$f_8 = x_9 \vee x_{10}$,
$f_9 = x_2 \wedge L_e \wedge \overline{G_e}$,
$f_{10} = ((L_{em} \wedge x_2) \vee L_e) \wedge \overline{G_e}$.

Lac operon Network



- Wiring diagram for the genes listed on the left. Arrows represent activation while blunt arrows inhibition.
- This system is responsible for the metabolism of lactose in the absence of glucose. This system exhibits bistability in the sense that the operon can be either ON or OFF, depending on the presence of the preferred energy source: glucose.

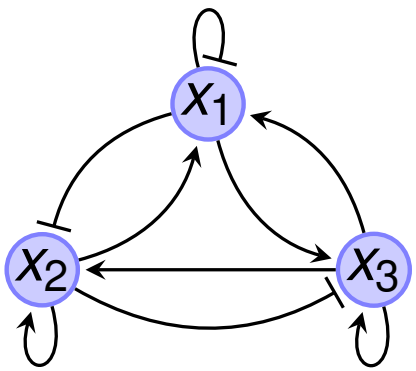
Boolean Models Can Explain Bistability in the lac Operon.
Alan Veliz Cuba and Brandilyn Stigler.
Journal of Computational Biology, 18:6, 2011.

Dynamical Systems over Finite Fields

$$\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$$

where

- $\mathbb{F}$ is a finite field of order $|\mathbb{F}| = q = p^m$.
- The functions $f_i : \mathbb{F}^n \to \mathbb{F}$ can be written as polynomials over the finite field $\mathbb{F}$. Thus **F** is a Polynomial Dynamical System.
- In particular, if $\mathbb{F}_2 = \{0, 1\}$ is the finite field with two elements, then **F** is a Boolean network.



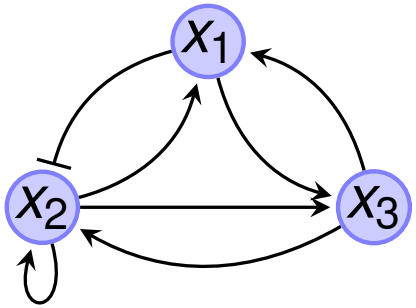
The AND rule: $x \wedge y = x \cdot y$

The OR rule: $x \vee y = x \cdot y + x + y$

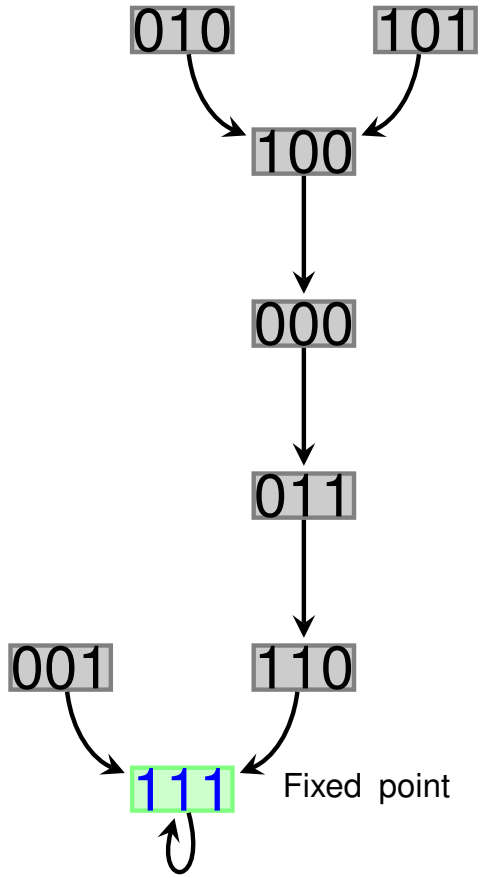
The negation: $\neg x = x + 1$

Example of a BN

Wiring Diagram



State Space

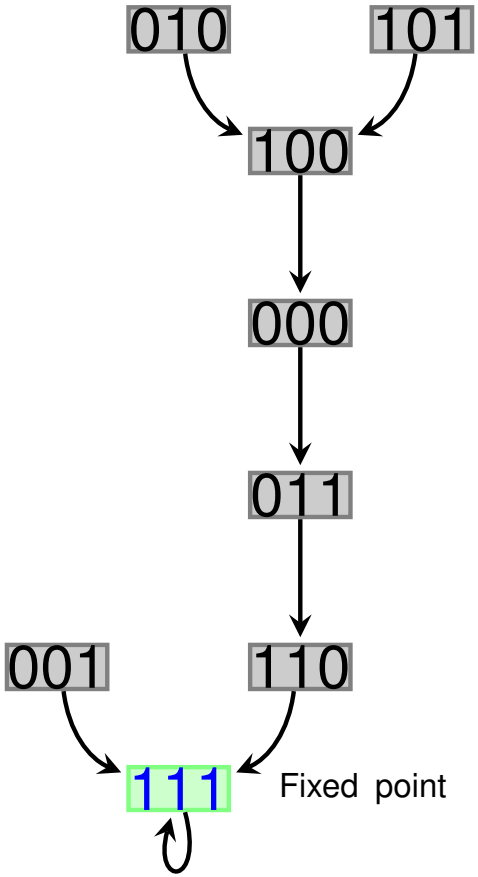


$$x(t+1) = \mathbf{F}(x(t))$$

where $\mathbf{F} = (f_1, f_2, f_3) : \mathbb{F}^3 \to \mathbb{F}^3$, $\mathbb{F} = \{0, 1\}$

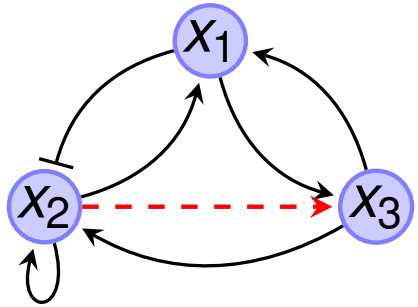
$$\begin{cases} f_1(x_2, x_3) = x_2x_3 + x_2 + x_3, \\ f_2(x_1, x_2, x_3) = x_1x_2x_3 + x_2x_3 + x_1 + x_2 + 1, \\ f_3(x_1, x_2) = x_1 + x_2 + 1. \end{cases}$$

Control in this setting

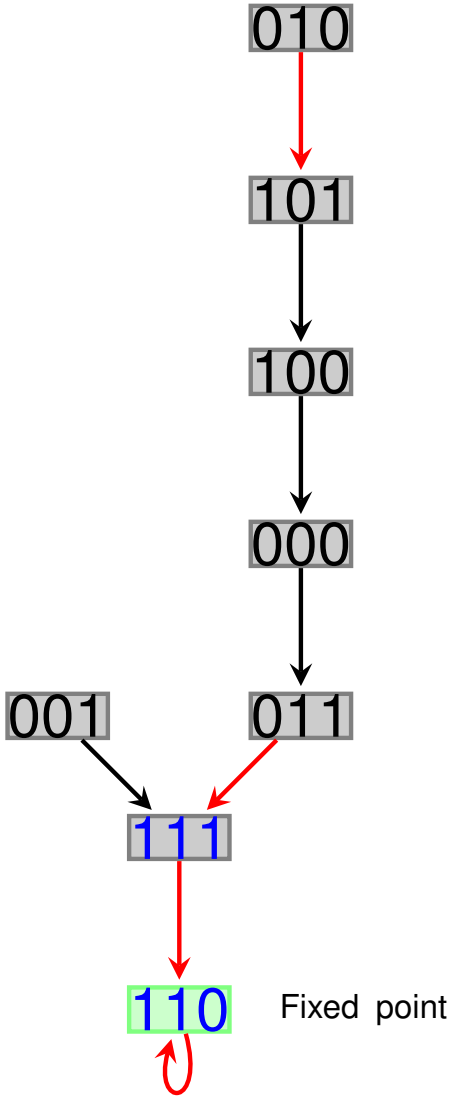


Original State Space.

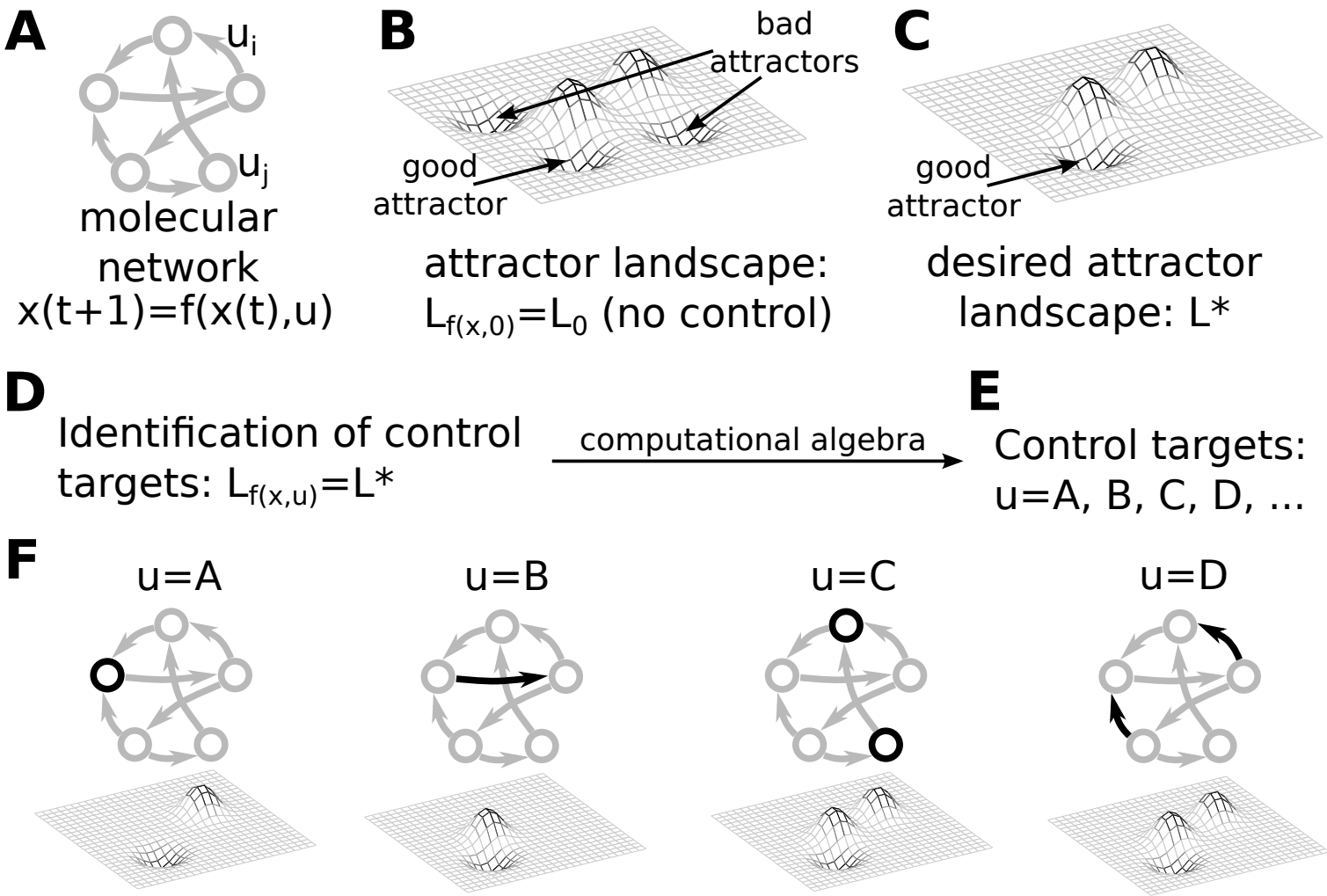
Wiring Diagram.



State space after edge deletion $\Rightarrow$



Network Control For Boolean networks

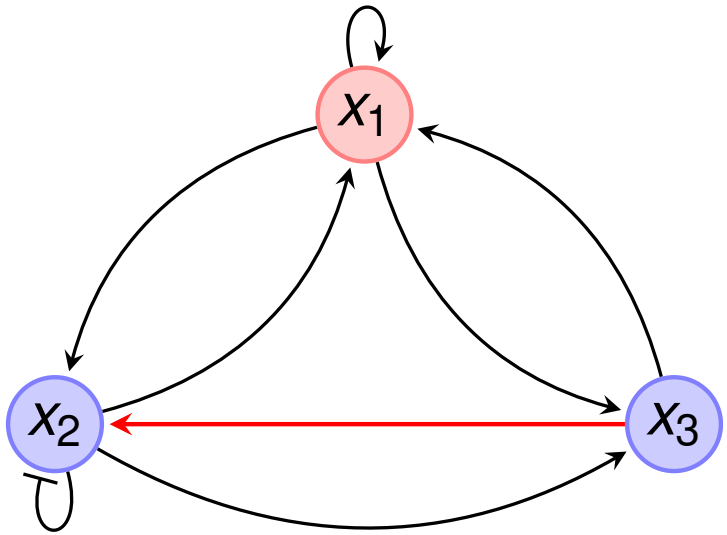


Identification of control targets in Boolean molecular network models via computational algebra.
D. Murrugarra, A. Veliz-Cuba, B. Aguilar, and R. Laubenbacher. BMC Systems Biology, 10:94, 2016.

Definition of Control Actions

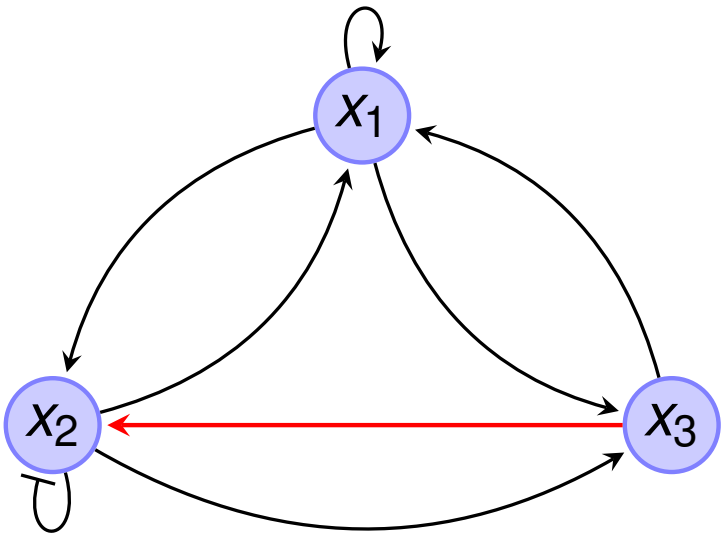
We consider two types of control actions:

1. Deletion or constant expression of edges
2. Deletion or constant expression of nodes.



Encoding Edge Deletions

Network



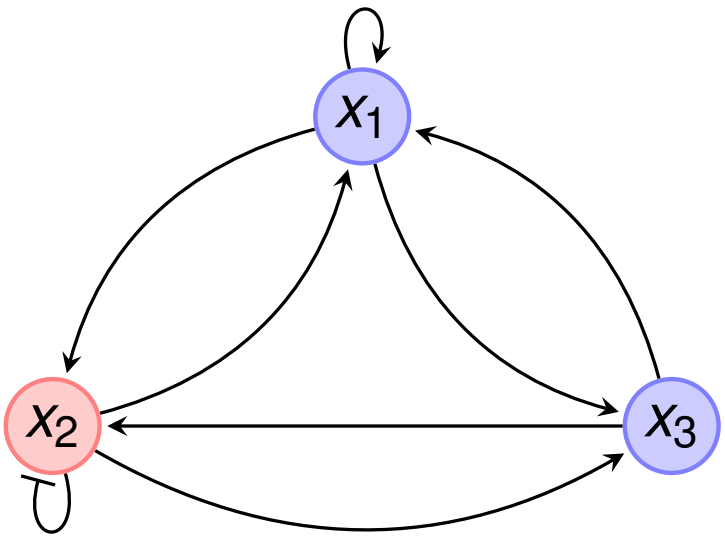
Controlled System

$$\mathcal{F}_2(\mathbf{x}, u_{3,2}) = f_2(x_1, x_2, (u_{3,2}+1)x_3)$$

- For $u_{3,2} = 0$, $\mathcal{F}_2(\mathbf{x}, 0) = f_2(x_1, x_2, x_3)$.
  The control is not active.
- For $u_{3,2} = 1$, $\mathcal{F}_2(\mathbf{x}, 1) = f_2(x_1, x_2, 0)$.
  The control is active and the action represents the deletion of the edge $x_3 \to x_2$.

Encoding Node Deletions

Network



Regulatory rule

$$\mathcal{F}_j(\mathbf{x}, u_i^-, u_i^+) := (u_i^- + u_i^+ + 1) f_j(\mathbf{x}) + u_i^+$$

- For $u_i^- = 0, u_i^+ = 0$, $\mathcal{F}_j(x, 0, 0) = f_j(x)$. The control is not active.
- For $u_i^- = 1, u_i^+ = 0$, $\mathcal{F}_j(x, 1, 0) = 0$. This action represents the knock out of the node x_j.
- For $u_i^- = 0, u_i^+ = 1$, $\mathcal{F}_j(x, 0, 1) = 1$. This action represents the constant expression of the node x_j.
- For $u_i^- = 1, u_i^+ = 1$, $\mathcal{F}_j(x, 1, 1) = f_j(x_{t_1}, \ldots, x_{t_m}) + 1$.

Generating new steady states

Let $\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ where $\mathbb{F} = \{0, 1\}$.

- Suppose that $\mathbf{y}_0 = (y_{01}, \ldots, y_{0n}) \in \mathbb{F}^n$ is a desirable cell state (for instance, it could represent the state of cell senescence).
- but $\mathbf{y}_0$ is not a fixed point, i.e., $\mathbf{F}(\mathbf{y}_0) \neq \mathbf{y}_0$.

Generating new steady states

Let $\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ where $\mathbb{F} = \{0, 1\}$.

- Suppose that $\mathbf{y}_0 = (y_{01}, \ldots, y_{0n}) \in \mathbb{F}^n$ is a desirable cell state (for instance, it could represent the state of cell senescence).
- but $\mathbf{y}_0$ is not a fixed point, i.e., $\mathbf{F}(\mathbf{y}_0) \neq \mathbf{y}_0$.

Goal

Find a set of controllers $\mu = \{\mu_1, \ldots, \mu_n\}$ so that $\mathcal{F}(\mathbf{y}_0, \mu) = \mathbf{y}_0$.

To solve this problem we consider the system of polynomial equations in the u parameters:

$$\mathcal{F}_j(\mathbf{y}_0, u) - y_{0j} = 0, j = 1, \ldots, m. \tag{1}$$

Destroying fixed points or blocking transitions

Given $\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ with $\mathbf{F}(\mathbf{x}_0) = \mathbf{x}_0$, for $\mathbf{x}_0 \in \mathbb{F}^n$.

Suppose that $\mathbf{x}_0$ is an undesirable attractor (it could represent a tumor proliferative cell state that needs to be avoided).

Destroying fixed points or blocking transitions

Given $\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ with $\mathbf{F}(\mathbf{x}_0) = \mathbf{x}_0$, for $\mathbf{x}_0 \in \mathbb{F}^n$.

Suppose that $\mathbf{x}_0$ is an undesirable attractor (it could represent a tumor proliferative cell state that needs to be avoided).

Goal

Find a set of control edges such that $\mathcal{F}(\mathbf{x}_0, \mu) \neq \mathbf{x}_0$.

To solve this problem consider the following equation,

$$[\mathcal{F}_1(\mathbf{x}, u_{j,1}) - x_{01} + 1] \cdots [\mathcal{F}_n(\mathbf{x}, u_{j,n}) - x_{0m} + 1] = 0 \tag{2}$$

In general, for blocking a transition, consider

$$[\mathcal{F}_1(\mathbf{x}, u_{j,1}) - z_{01} + 1] \cdots [\mathcal{F}_n(\mathbf{x}, u_{j,n}) - z_{0n} + 1] = 0 \tag{3}$$

Blocking regions in the state space

Let $\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ where $\mathbb{F} = \{0, 1\}$.

Suppose a particular value of a variable, $x_k = a \in \mathbb{F}_2$, triggers an undesirable pathway, or is the signature of an abnormal cell, then we want all steady states of the system to satisfy $x_k \neq a$.

Blocking regions in the state space

Let $\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ where $\mathbb{F} = \{0, 1\}$.

Suppose a particular value of a variable, $x_k = a \in \mathbb{F}_2$, triggers an undesirable pathway, or is the signature of an abnormal cell, then we want all steady states of the system to satisfy $x_k \neq a$.

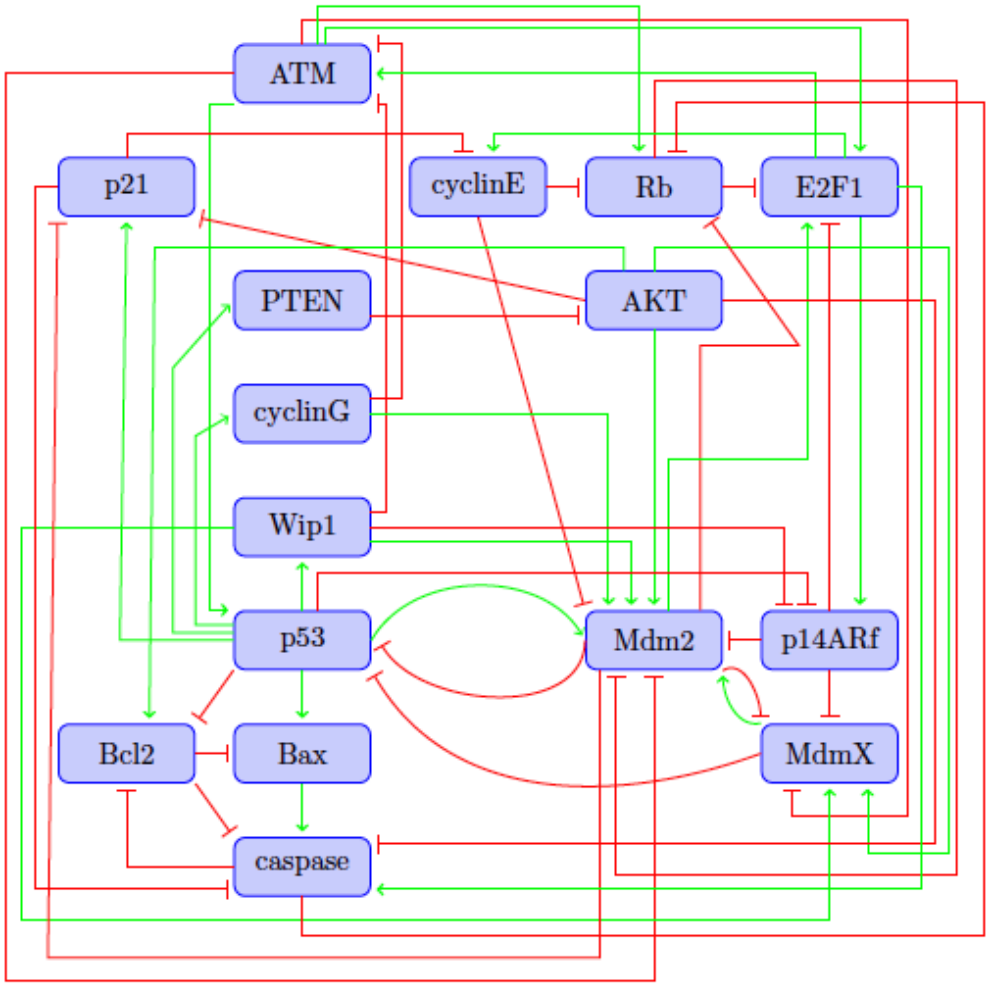
Goal

In this case, we consider the systems of equations

$$\begin{aligned} \mathcal{F}_j(x, u) - x_j = 0, j = 1, \ldots, m, \\ x_k - a = 0. \end{aligned} \tag{4}$$

Since the steady states with $x_k = a$ are to be avoided, we want to find controls u for which Equation 4 has no solution.

Application: The p53-mdm2 complex



Wiring diagram for the p53-mdm2 complex. Adapted from Choi et al. Science Signaling 5 (251), 2012.

We will use the cancer cell model where *PTEN* and *p14ARf* are always inactive (fixed to zero) and *cyclinG* is always active (fixed to 1).

This is a Boolean network $\mathbf{F} = (f_1, \ldots, f_{16}) : \mathbb{F}_2^{16} \to \mathbb{F}_2^{16}$ with 16 nodes and 50 edges. We represent the nodes by

$x_1 = ATM$, $x_2 = p53$,
$x_3 = Mdm2$, $x_4 = MdmX$,
$x_5 = Wip1$, $x_6 = cyclinG$,
$x_7 = PTEN$, $x_8 = \text{p21}$,
$x_9 = AKT$, $x_{10} = cyclinE$,
$x_{11} = Rb$, $x_{12} = E2F1$,
$x_{13} = p14ARf$, $x_{14} = Bcl2$,
$x_{15} = Bax$, $x_{16} = caspase$.

Identification of control targets in Boolean molecular network models via computational algebra.
David Murrugarra, Alan Veliz-Cuba, Boris Aguilar, and Reinhard Laubenbacher.
BMC Systems Biology, 10:94, 2016.

PDS for p53-mdm2 complex

$$\mathbf{f_1}(x_1, x_5, x_6, x_{12}) = x_1x_{12} + x_1x_{12}x_5 + x_1x_{12}x_6 + x_1x_{12}x_5x_6,$$

$$\mathbf{f_2}(x_1, x_2, x_3, x_4) = 1 + x_3 + x_1x_2x_3 + x_4 + x_1x_4 + x_3x_4 + x_1x_3x_4 + x_1x_2x_3x_4,$$

$$\mathbf{f_3}(x_1, x_2, x_3, x_4, x_5, x_6, x_9, x_{10}, x_{11}, x_{13}) = 1 + x_1 + x_{10} + x_1x_{10} + x_{11} + x_1x_{11} + x_{10}x_{11} + x_1x_{10}x_{11}$$
$$+x_{13} + x_1x_{13} + x_{10}x_{13} + x_1x_{10}x_{13} + x_{11}x_{13} + x_1x_{11}x_{13} + x_{10}x_{11}x_{13} + x_1x_{10}x_{11}x_{13} + x_1x_2x_3$$
$$+x_{10}x_2x_3 + x_{11}x_2x_3 + x_1x_{10}x_{11}x_2x_3 + x_1x_{13}x_2x_3 + x_{10}x_{13}x_2x_3 + x_{11}x_{13}x_2x_3 + x_1x_{10}x_{11}x_{13}x_2x_3$$
$$+x_1x_4 + x_{10}x_4 + x_{11}x_4 + x_1x_{10}x_{11}x_4 + x_1x_{13}x_4 + x_{10}x_{13}x_4 + x_{11}x_{13}x_4 + x_1x_{10}x_{11}x_{13}x_4$$
$$+x_{13}x_2x_4 + x_1x_{13}x_2x_4 + x_{10}x_{13}x_2x_4 + x_1x_{10}x_{13}x_2x_4 + x_{11}x_{13}x_2x_4 + x_1x_{11}x_{13}x_2x_4 + x_{10}x_{11}x_{13}x_2x_4$$
$$+x_1x_{10}x_{11}x_{13}x_2x_4 + x_{13}x_3x_4 + x_1x_{13}x_3x_4 + x_{10}x_{13}x_3x_4 + x_1x_{10}x_{13}x_3x_4 + x_{11}x_{13}x_3x_4$$
$$+x_1x_{11}x_{13}x_3x_4 + x_{10}x_{11}x_{13}x_3x_4 + x_1x_{10}x_{11}x_{13}x_3x_4 + x_1x_2x_3x_4 + x_{10}x_2x_3x_4 + x_1x_{10}x_2x_3x_4$$
$$+x_{11}x_2x_3x_4 + x_1x_{11}x_2x_3x_4 + x_{10}x_{11}x_2x_3x_4 + x_{13}x_2x_3x_4 + x_1x_{10}x_{11}x_{13}x_2x_3x_4 + x_1x_5 + x_{10}x_5$$
$$+x_{11}x_5 + x_1x_{10}x_{11}x_5 + x_1x_{13}x_5 + x_{10}x_{13}x_5 + x_{11}x_{13}x_5 + x_1x_{10}x_{11}x_{13}x_5 + x_{13}x_2x_5 + x_1x_{13}x_2x_5$$
$$+x_{10}x_{13}x_2x_5 + x_1x_{10}x_{13}x_2x_5 + x_{11}x_{13}x_2x_5 + x_1x_{11}x_{13}x_2x_5 + x_{10}x_{11}x_{13}x_2x_5 \ldots$$

Polynomial functions for the p53 network.

Attractor for the p53-mdm2 complex

In the presence of DNA damage this network has a single limit cycle of size 7 representing the state of cell cycle arrest, where *p*53 and *p*21 are oscillating.

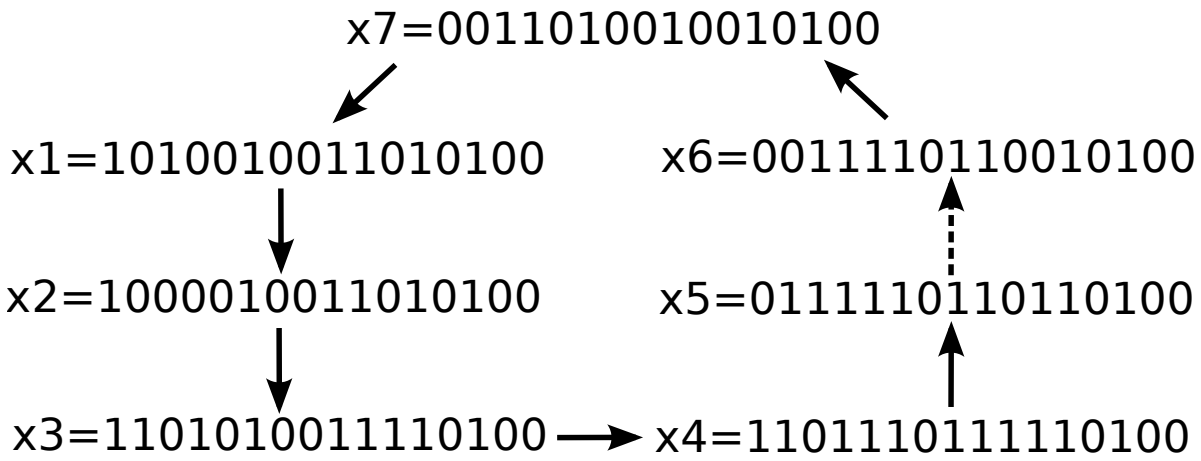


Figure: States of limit cycle representing cell cycle arrest in the p53 model. The dashed edge represents the transition target to destroy the limit cycle.

Identification of control edges

Goal

1. Make $\mathbf{y}_0$ a fixed point of the system.
2. Destroy this limit cycle.

In this case, let us consider the state that represents the cell death,

$$\mathbf{y}_0 = (1,1,1,0,0,1,0,1,1,0,0,1,0,0,1,1), \text{ where } x_{16} = \textit{caspase} \text{ is active.}$$

In order to make $\mathbf{y}_0$ a fixed point, we consider the following system of polynomial equations,

$$\mathcal{F}_i(\mathbf{y}_0, \mu) = y_{0i}, j = 1, \ldots, 16. \tag{5}$$

The solutions for the system of equations 5 are given by the nonzero generators of the ideal associated to the system 5, given in Equation 6,

$$\begin{aligned}&\{u_{2,5}+1, u_{16,11}(u_{1,11}+1), u_{8,8}(u_{3,8}+1), u_{2,2}(u_{3,2}+1),\\ &u_{1,2}(u_{3,2}+1), u_{1,1}u_{12,1}, u_{12,16}u_{16,16}(u_{8,16}+1)\}\end{aligned} \tag{6}$$

Identification of control edges

Controllers applied	Ref.	Basin size of $\mathbf{y}_0$
$mdm2 \to p53$ $p53 \to Wip1$	Choi et al. 2012.	35581 (54.29%)
$p53 \to Wip1$ $Mdm2 \to p21$ $Mdm2 \to p53$ $p21 \to Caspase$	A control set that forces $\mathbf{y}_0$ to be a fixed point.	39856 (60.82%)
$mdm2 \to p53$ $p53 \to Wip1$ $mdm2 \to p21$ $p21 \to Caspase$ $ATM \to Rb$ $mdm2 \to Rb$ $mdmx \to p53$ $Rb \to E2F1$ $Bax \to Caspase$	A control set to make $\mathbf{y}_0$ a fixed point and for blocking the dashed transition.	65536 (100%)

Effect of an edge deletion on the state space

- What are the side effects of applying a given control?
- What is the effect of an edge deletion in the state space?

To answer this question we count the maximum number of state space transitions that can be changed as a result of deleting a single edge.

Boolean Canalizing Rules

$$f(x_1, x_2, x_3) = (x_1 + 1)[x_2(x_3 + 1) + 1] + 1$$

$$f(x_1 = 1, x_2, x_3) = 1 \text{ (the variable } x_1 \text{ is canalizing).}$$

The variables x_2 and x_3 are not canalizing.

$$\begin{aligned} f(x_1, x_2 = 0, x_3) &= x_1, \\ f(x_1, x_2 = 1, x_3) &= (x_1 + 1)x_3 + 1, \\ f(x_1, x_2, x_3 = 0) &= (x_1 + 1)(x_2 + 1) + 1, \\ f(x_1, x_2, x_3 = 1) &= x_1. \end{aligned}$$

S. Kauffman introduced the concept of canalizing Boolean rules.
The Origins of Order: Self-Organization and Selection in Evolution. S. A. Kauffman. *Oxford University Press, New York, Oxford*, 1993.

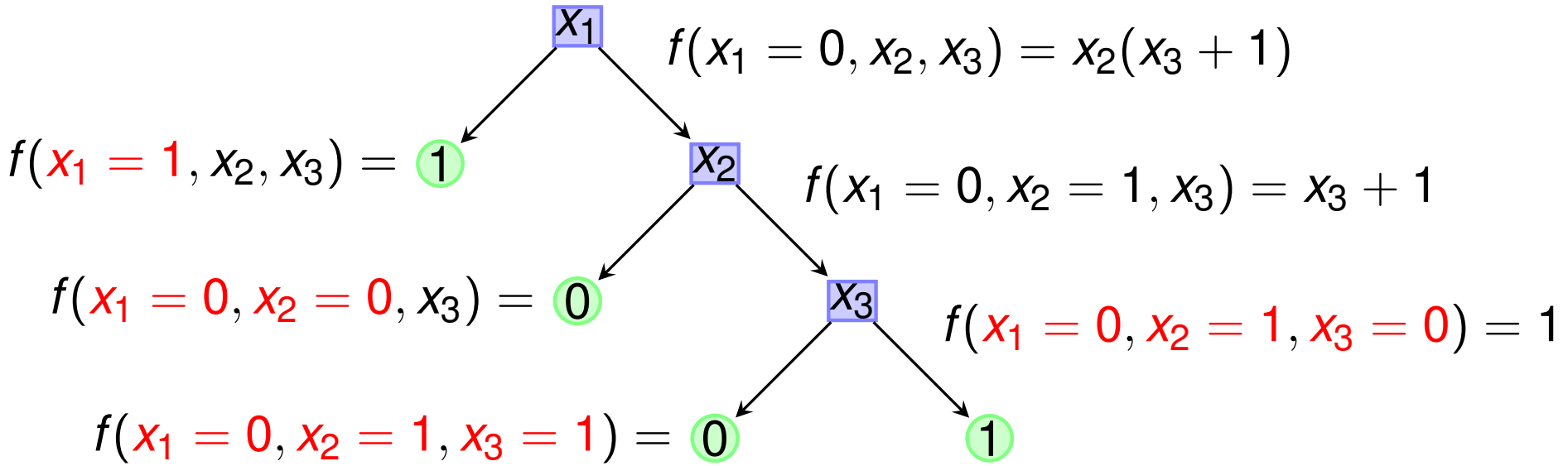
The term *canalization* was coined by the geneticist C.H. Waddington in the 1940's.
Canalization of Development and the Inheritance of Acquired Characters. C.H. Waddington. *Nature.* 150:563-565, 1942.

Nested Canalizing Functions

The function

$$f(x_1, x_2, x_3) = (x_1 + 1)[x_2(x_3 + 1) + 1] + 1$$

is *nested canalizing* in the variable order x_1, x_2, x_3 with *canalizing input values* $1, 0, 1 \in \mathbb{F}_2$ and *canalized output values* $1, 0, 0, 1 \in \mathbb{F}_2$.



Layers of canalization for Boolean networks

For the previous function

$$\begin{aligned} f_2(x_1, x_2, x_3) &= (x_1 + 1)[x_2(x_3 + 1) + 1] + 1 \\ &= M_1[M_2 + 1] + 1 \end{aligned}$$

1. $M_1 = x_1 + 1$ is the first layer.
2. $M_2 = x_2(x_3 + 1)$ is the second layer.

- Boolean canalization gives a hierarchical clustering of the variables. For instance, x_1 is the most dominant variable.
- A Boolean function can be represented in different forms as a nested canalizing function.
- A unique representation of the function is obtained by grouping the variables in layers of canalization.

Boolean nested canalizing functions: a comprehensive analysis.
Y. Li, J. O. Adeyeye, D. Murrugarra, B. Aguilar, R. Laubenbacher.
Theoretical Computer Science, 481, 24-36, 2013.

Layer number of a nested canalizing function

Polynomial form

Let $f : \mathbb{F}^n \to \mathbb{F}$ be a nested canalyzing function. Then f can uniquely be written as a nested product of extended monomials, i.e.

$$f(x_1, \ldots, x_n) = M_1 \left[M_2(\ldots(M_m + 1)\ldots) + 1 \right] + b_1$$

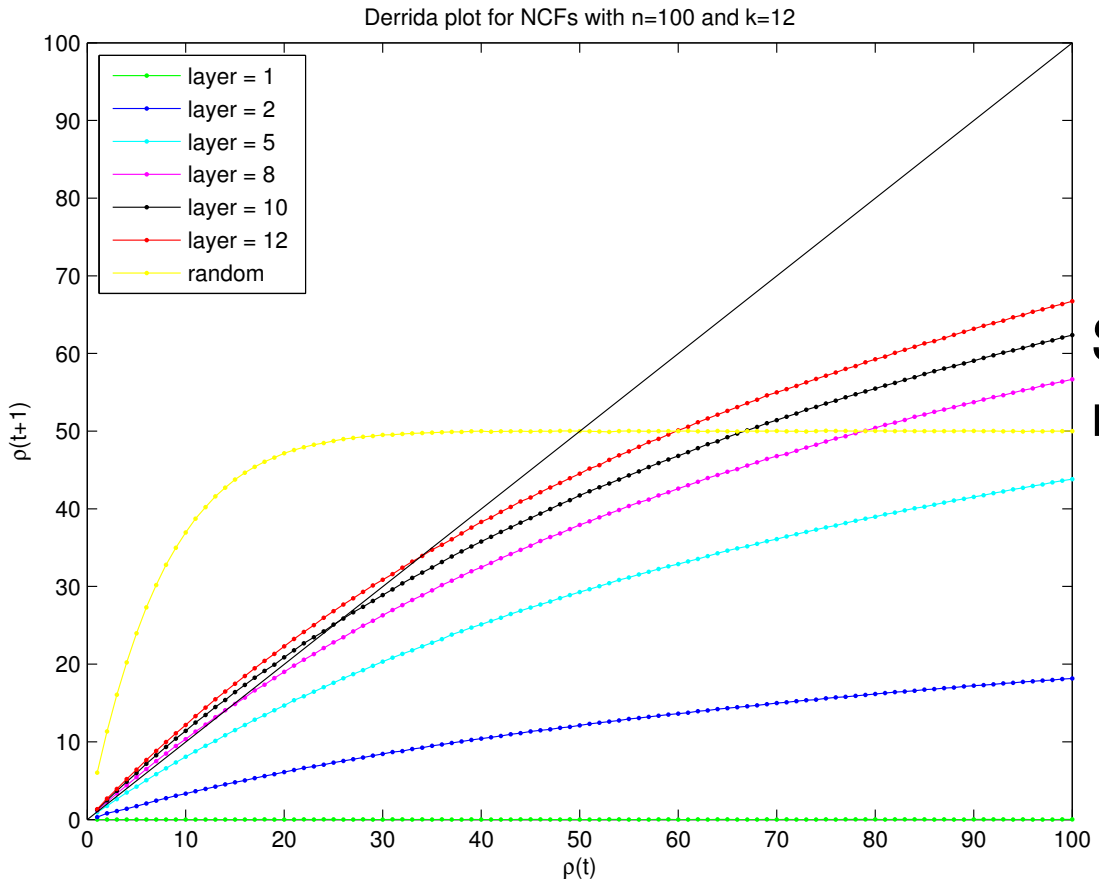
$$\text{where } M_i = \prod_{i=k_{i-1}+1}^{k_i} (x_i - a_i)$$

for $i = 1, \ldots, m$.

Layer number of f: m

Boolean nested canalizing functions: a comprehensive analysis.
Y. Li, J. O. Adeyeye, D. Murrugarra, B. Aguilar, R. Laubenbacher.
Theoretical Computer Science, 481, 24-36, 2013.

Finer categorization of NCFs



Stability of networks with nested canalizing functions.

Boolean nested canalizing functions: a comprehensive analysis.
Y. Li, J. O. Adeyeye, D. Murrugarra, B. Aguilar, R. Laubenbacher.
Theoretical Computer Science, 481, 24-36, 2013.

Stratification of variables for Boolean networks

Theorem

Every Boolean function can be uniquely written as

$$f(x_1, \ldots, x_n) = M_1(M_2(\ldots(M_{m-1}(M_m P_c + 1) + 1)\ldots) + 1) + b, \quad (7)$$

where $M_i = \prod_{i=1}^{k_i}(x_{i_j} + a_{i_j})$, P_c is a polynomial with no canalizing variables, and $k = k_1 + \cdots + k_m$ is the canalizing depth. Each variable x_i appears in exactly one of the $M_1, M_2, \ldots, M_m, P_c$. The number m in Equation 7 is called the *layer number* of f.

Stratification and enumeration of Boolean functions by canalizing depth.
Qijun He and Matthew Macauley.
Physica D: Nonlinear Phenomena, 314, 1-8, 2015.

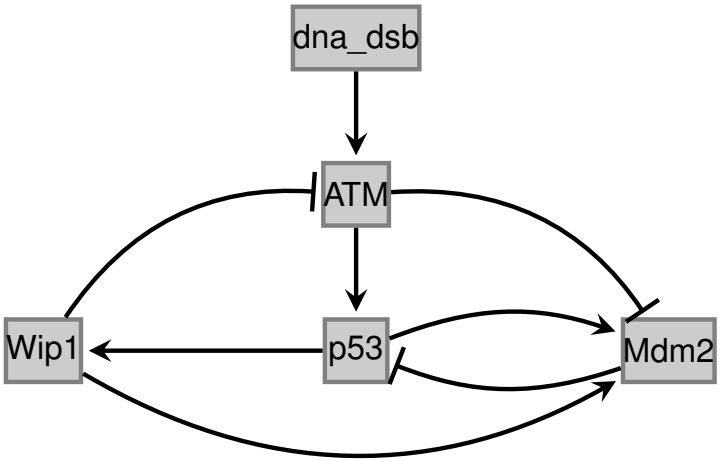
Examples

Consider the Boolean functions

1. $f_1(x_1, x_2, x_3) = x_1x_3(x_2 + 1) + 1$, where $M_1 = x_1x_3(x_2 + 1)$.
   Thus f_1 has layer number equal to 1.
2. $f_2(x_1, x_2, x_3) = (x_1 + 1)[x_2(x_3 + 1) + 1] + 1$, where
   $M_1 = x_1 + 1$, $M_2 = x_2(x_3 + 1)$.
   Thus f_2 has layer number equal to 2.
3. $f_3(x_1, x_2, x_3) = (x_1 + 1)(x_2 + x_3 + 1)$, where $M_1 = x_1 + 1$,
   $P_c = x_2 + x_3 + 1$.
   Thus f_3 has layer number 1. The core polynomial P_c does not have canalizing variables.

Application: *p53-mdm2* network

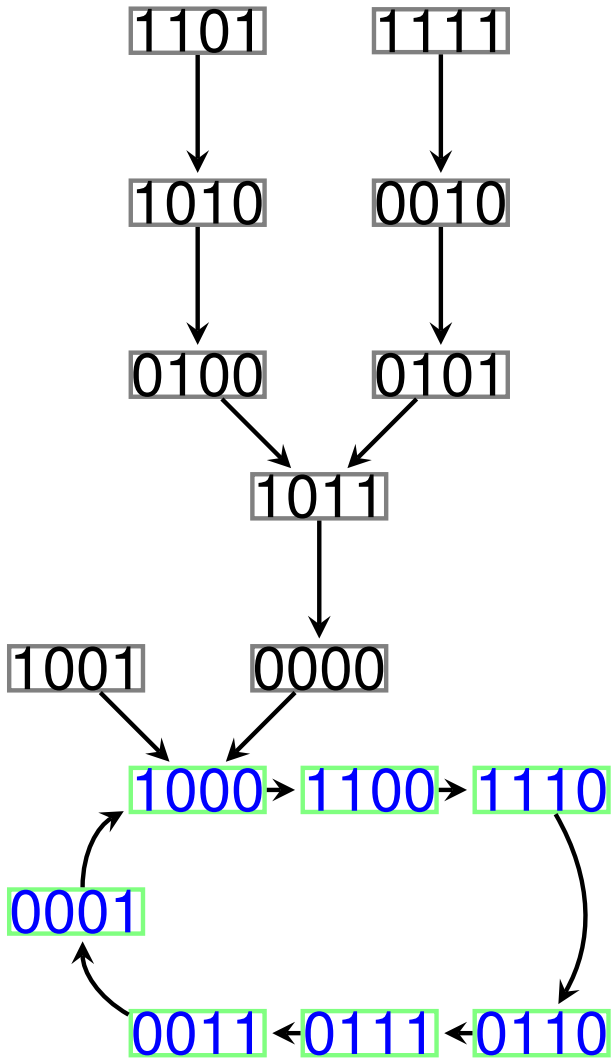
Wiring diagram of the *p53-mdm2* model.



Regulatory rules.

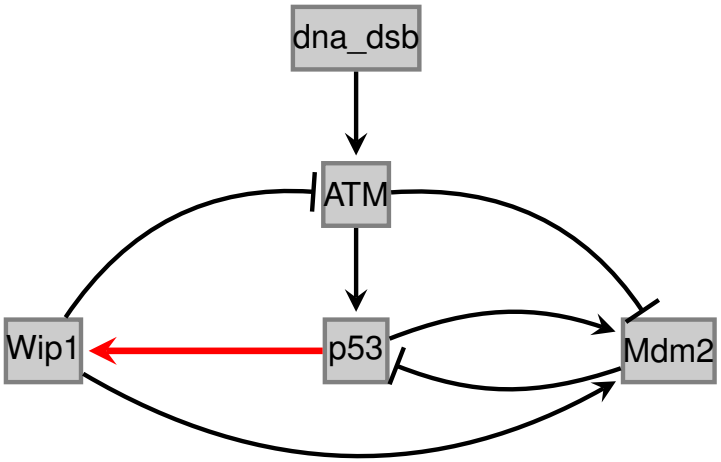
$$
\begin{aligned}
\mathrm{ATM}_{\mathrm{next}} &= \overline{\mathrm{Wip1}}(\mathrm{ATM} + \mathrm{dna\_dsb}) \\
\mathrm{p53}_{\mathrm{next}} &= \overline{\mathrm{Mdm2}}(\mathrm{ATM} + \mathrm{Wip1}) \\
\mathrm{Wip1}_{\mathrm{next}} &= \mathrm{p53} \\
\mathrm{Mdm2}_{\mathrm{next}} &= \overline{\mathrm{ATM}}(\mathrm{p53} + \mathrm{Wip1})
\end{aligned}
$$

State Space

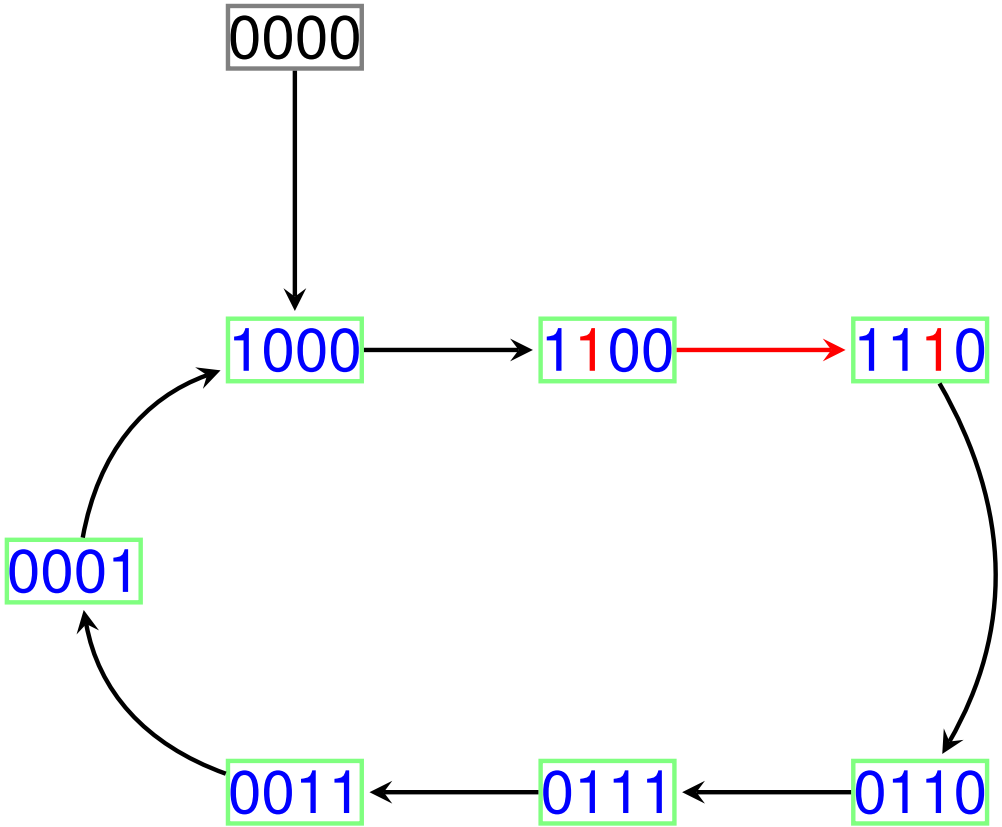


Application: *p53-mdm2* network

Wiring diagram of the *p53-mdm2* model.



State Space

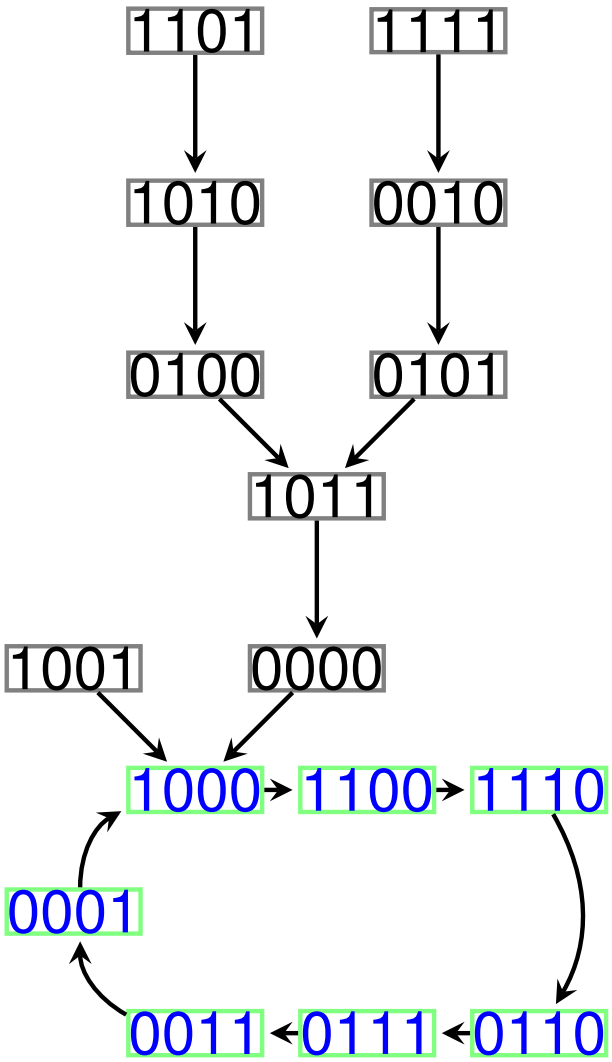


Regulatory rules.

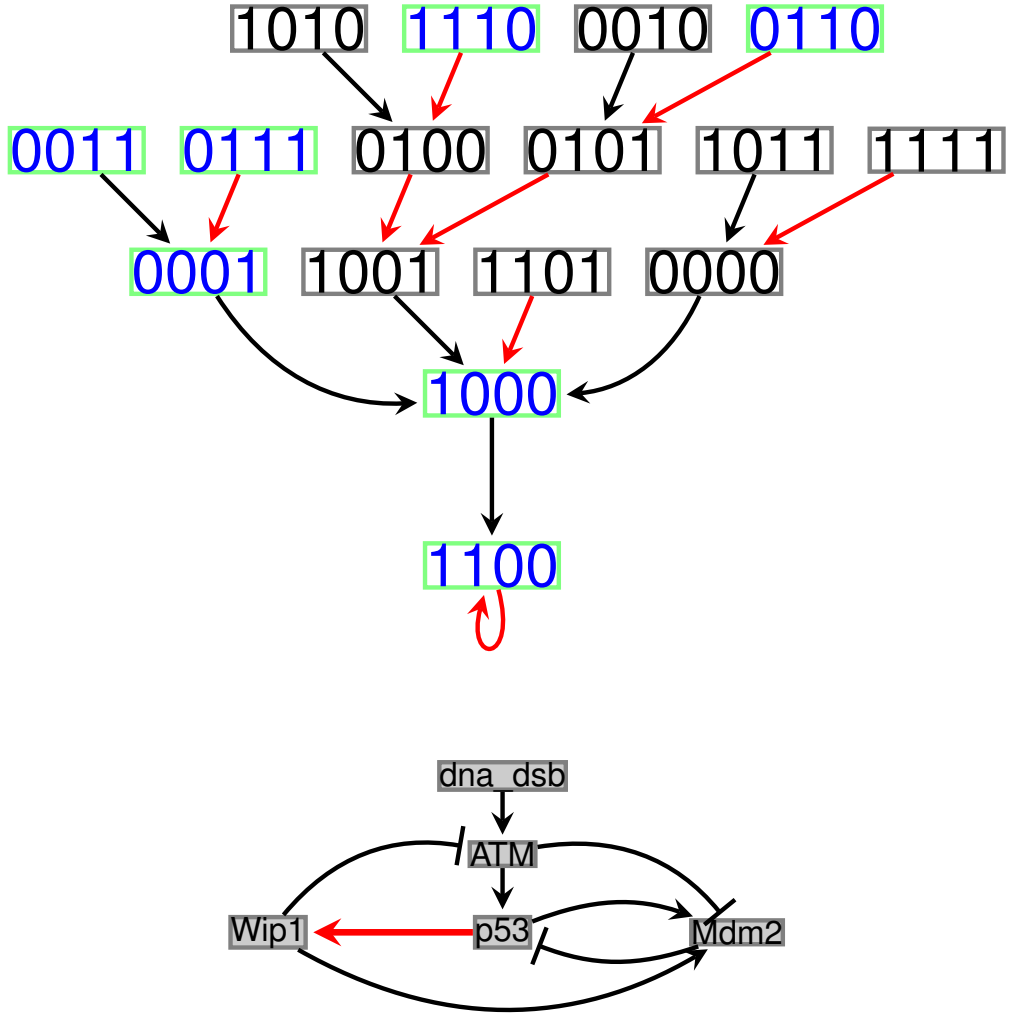
$$
\begin{aligned}
\mathrm{ATM}_{\mathrm{next}} &= \overline{\mathrm{Wip1}}(\mathrm{ATM} + \mathrm{dna\_dsb}) \\
\mathrm{p53}_{\mathrm{next}} &= \overline{\mathrm{Mdm2}}(\mathrm{ATM} + \mathrm{Wip1}) \\
\mathrm{Wip1}_{\mathrm{next}} &= \mathrm{p53} \\
\mathrm{Mdm2}_{\mathrm{next}} &= \overline{\mathrm{ATM}}(\mathrm{p53} + \mathrm{Wip1})
\end{aligned}
$$

Application: *p53-mdm2* network

State Space (No Control).



State Space (Under Control).



Effect of an edge deletion on the state space

Theorem (Upper Bound)

Let $\mathbf{F} = (f_1, \ldots, f_n) : \{0,1\}^n \to \{0,1\}^n$ *be a Boolean network where*

$$f_t(x_1, \ldots, x_n) = M_1(M_2(\ldots(M_{m-1}(M_m P_c + 1) + 1)\ldots) + 1) + b,$$

where $M_i = \prod_{i=1}^{\ell_i}(x_{i_j} + a_{i_j})$, P_c *is a polynomial with no canalizing variables, and* $d = \ell_1 + \ell_2 + \cdots + \ell_m$ *is the canalizing depth.*
The probability that any transition will be removed from the state space upon deletion of $x_k \to x_t$ *is at most*

$$2^{n-\ell_1-\ell_2-\ldots-\ell_r}/2^n = \left(\frac{1}{2}\right)^{\ell_1+\ell_2+\cdots+\ell_r}.$$

Molecular Network Control Through Boolean Canalization.
David Murrugarra and Elena Dimitrova.
EURASIP Journal on Bioinformatics and Systems Biology, 2015:9, 2015.

Application: *p53-mdm2* network

Example

For this example, $n = 4$, $r = 1$, $\ell_1 = 1$, and

$$\text{Wip1}_{\text{next}} = \text{p53}$$

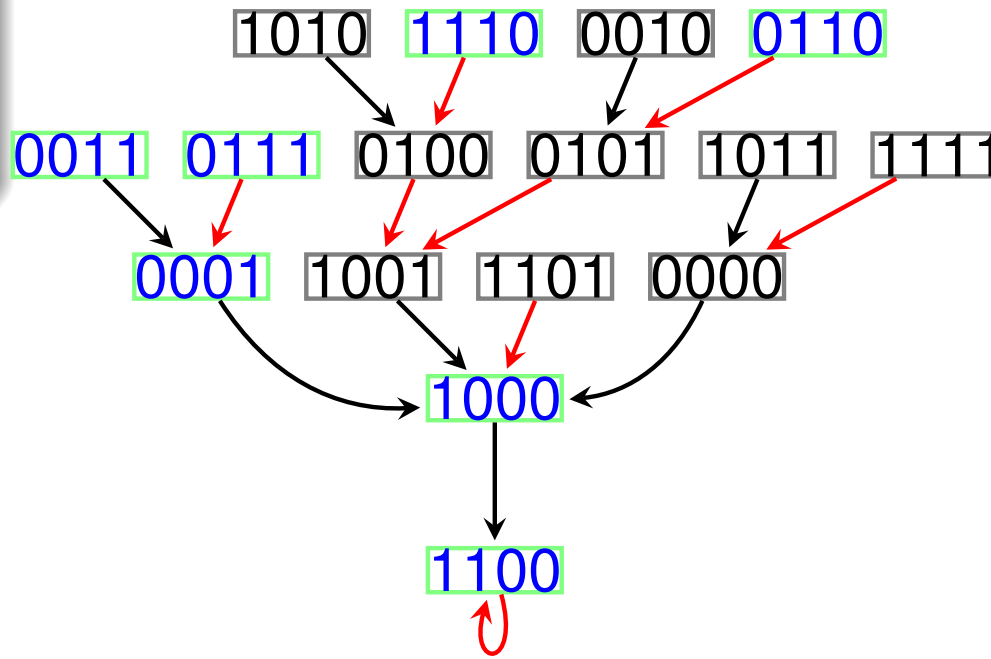
The deletion of the edge *p53* $\rightarrow$ *Wip1* results in up to

$$2^{n-\ell_1} = 2^{4-1} = 8.$$

changes in the state space.

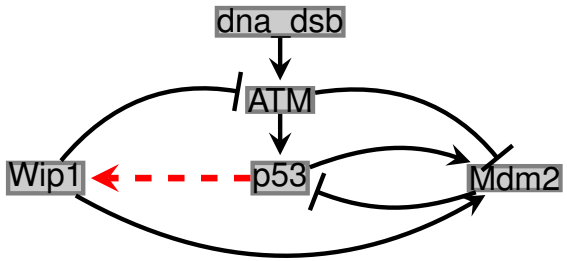
State Space

State Space (Under Control).



Regulatory rules.

$$\begin{aligned}
\text{ATM}_{\text{next}} &= \overline{\text{Wip1}}(\text{ATM} + \text{dna_dsb}) \\
\text{p53}_{\text{next}} &= \overline{\text{Mdm2}}(\text{ATM} + \text{Wip1}) \\
\text{Wip1}_{\text{next}} &= \text{p53} \\
\text{Mdm2}_{\text{next}} &= \overline{\text{ATM}}(\text{p53} + \text{Wip1})
\end{aligned}$$

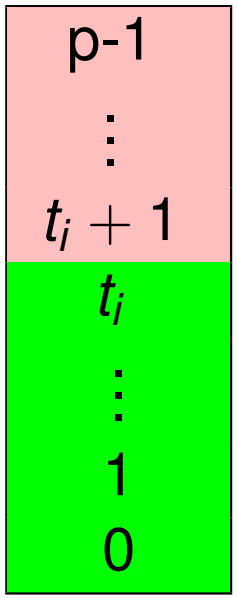


Nested Canalizing Functions: multistate case

Let $\mathbb{F}$ be a finite field with p elements. Let's arrange its elements in increasing order by $\mathbb{F} = \{0 < 1 < \cdots < p-1\}$. Let $x_1, \ldots, x_n$ be variables which can take values in $\mathbb{F}$.

Let $S_i \subset \mathbb{F}$, $i = 1, \ldots, n$, such that

- S_i is a subinterval of $\mathbb{F}$.
- The complement of S_i is also a subinterval.



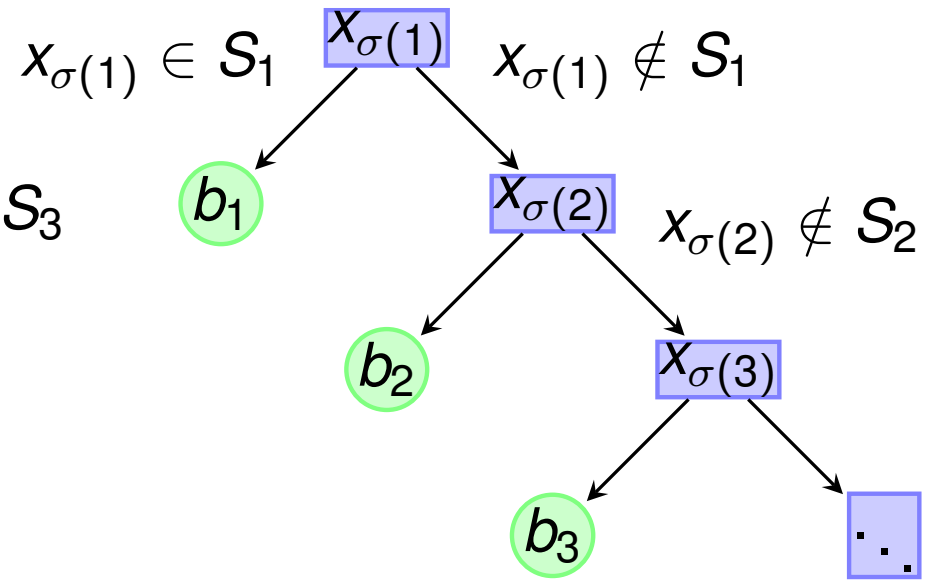
Regulatory Patterns in Molecular Interaction Networks.
D. Murrugarra and R. Laubenbacher.
Journal of Theoretical Biology (288), 66-72, 2011.

Nested Canalizing Functions: multistate case

Let σ be a permutation on $\{1,\ldots,n\}$. The function $f:\mathbb{F}^n\to\mathbb{F}$ is a *nested canalizing function* in the variable order $x_{\sigma(1)},\ldots,x_{\sigma(n)}$ with *canalizing input sets* $S_1,\ldots,S_n\subset\mathbb{F}$ and *canalized output values* $b_1,\ldots,b_n,b_{n+1}\in\mathbb{F}$ with $b_n\neq b_{n+1}$ if it can be represented in the form:

$$f(x_1,\ldots,x_n)=$$

$$\begin{cases} b_1 \text{ if } x_{\sigma(1)}\in S_1 \\ b_2 \text{ if } x_{\sigma(1)}\notin S_1, x_{\sigma(2)}\in S_2 \\ b_3 \text{ if } x_{\sigma(1)}\notin S_1, x_{\sigma(2)}\notin S_2, x_{\sigma(3)}\in S_3 \\ \vdots \\ b_n \text{ if } x_{\sigma(1)}\notin S_1,\ldots,x_{\sigma(n)}\in S_n \\ b_{n+1} \text{ if } x_{\sigma(1)}\notin S_1,\ldots,x_{\sigma(n)}\notin S_n \end{cases}$$

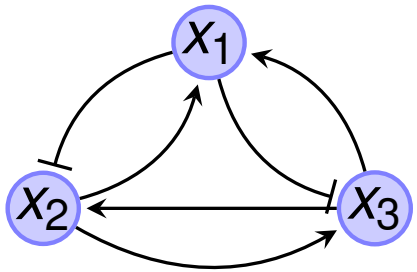


Regulatory Patterns in Molecular Interaction Networks.
D. Murrugarra and R. Laubenbacher.
Journal of Theoretical Biology (288), 66-72, 2011.

Dynamic Properties: stability

Networks with nested canalizing rules exhibit more stable dynamics compared to random networks.

Structure:



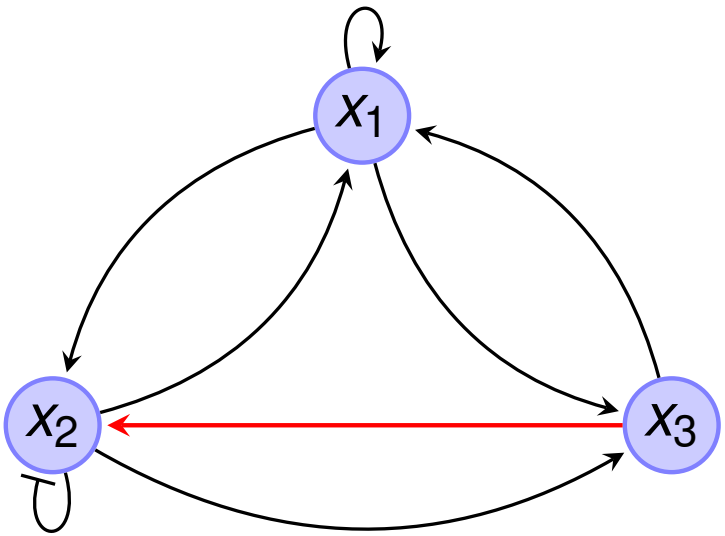
$$\begin{cases} x_1(t+1) = f_1(x_2(t), x_3(t)) \\ x_2(t+1) = f_2(x_1(t), x_3(t)) \\ x_3(t+1) = f_3(x_1(t), x_2(t)) \end{cases}$$

Dynamics.

- Nested canalyzing functions.
  - Few attractors, large basins.
  - Short limit cycles.
- Random function.
  - Many attractors, small basins.
  - Long limit cycles.

Encoding Edge Deletions: multistate case

Network



Controlled System

For $S_{3,2} \subset \mathbb{F}$,

$$\mathcal{F}_2(\mathbf{x}, Q_{S_{3,2}}(u)) = f_2(x_1, x_2, Q_{S_{3,2}}(u)x_3)$$

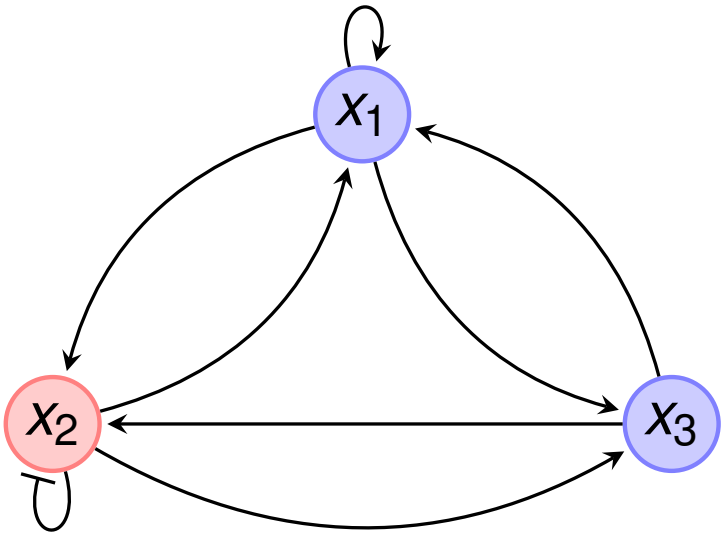
where

$$Q_{S_{3,2}}(u_0) = \begin{cases} 1 & \text{if } u_0 \in S_{3,2}, \\ 0 & \text{if } u_0 \notin S_{3,2}. \end{cases}$$

- For $u \in S_{3,2}$, $\mathcal{F}_2(\mathbf{x}, 1) = f_2(x_1, x_2, x_3)$.
  The control is not active.
- For $u \notin S_{3,2}$, $\mathcal{F}_2(\mathbf{x}, 0) = f_2(x_1, x_2, 0)$.
  The control is active and the action represents the deletion of the edge $x_3 \to x_2$.

Encoding Node Deletions: multistate case

Network



Regulatory rule

For $S_2 \subset \mathbb{F}$,

$$\mathcal{F}_2(\mathbf{x}, Q_{S_2}(u)) := Q_{S_2}(u) f_2(\mathbf{x})$$

where

$$Q_{S_2}(u_0) = \begin{cases} 1 & \text{if } u_0 \in S_2, \\ 0 & \text{if } u_0 \notin S_2. \end{cases}$$

- For $u \in S_2$, $\mathcal{F}_2(x, 1) = f_2(x)$. The control is not active.
- For $u \notin S_2$, $\mathcal{F}_2(x, 0) = 0$. This action represents the knock out of the node x_2.

Layers of Canalization for Multistate Functions

Theorem

Every multistate nested canalizing function can be uniquely written as

$$f(x_1,\dots,x_n) = M_1(M_2(\dots(M_{m-1}(B_{m+1}M_m+B_m)+B_{m-1})\dots)+B_2)+B_1,$$

where $M_i = \prod_{i=1}^{k_i} Q_{S_{i_j}}$, $n = k_1 + \cdots + k_m$, $B_1, B_2, \dots, B_{m+1} \in \mathbb{F}$, and $B_{m+1} \neq 0$.
Each variable x_i appears in exactly one of the $M_1, M_2, \dots, M_m$.

Definition: The layer number of $f(x_1,\dots,x_n)$ is m

Multistate nested canalizing functions and their networks.
C. Kadelka, Y. Li, J. Kuipers, J. O. Adeyeye, R. Laubenbacher.
Theoretical Computer Science, 675, 1-14, 2017.

Effect of an edge deletion on the state space.

Theorem (M.)

Let $\mathbf{F} = (f_1, \ldots, f_n) : \mathbb{F}^n \to \mathbb{F}^n$ *be a NCF network where*

$$f_t(x_1, \ldots, x_n) = M_1(M_2(\ldots(M_{m-1}(B_{m+1}M_m + B_m) + B_{m-1})\ldots) + B_2) + B_1,$$

where $M_i = \prod_{i=1}^{k_i} Q_{S_{i_j}}$, $n = k_1 + \cdots + k_m$, $B_1, \ldots, B_{m+1} \in \mathbb{F}$, *and* $B_{m+1} \neq 0$. *Each variable* x_i *appears in exactly one of the* $M_1, \ldots, M_m$.

If x_k *is in the* r^{th} *layer, then the probability that any transition will be removed from the state space upon deletion of* $x_k \to x_t$ *is at most*

$$|S_k|\left(p^{n-1} - \sum_{j=1}^{\ell_1-\ell_2-\cdots-\ell_r} |S_j|p^{n-j}\right)/p^n = \frac{|S_k|}{p} - \sum_{j=1}^{\ell_1+\ell_2+\cdots+\ell_r} \frac{|S_j|}{p^j}.$$

Stratification of Variables for Multistate Functions

Problem: Is there a stratification of variables by canalizing depth for multistate functions?

Conjecture

Every multistate function can be uniquely written as

$$f(x_1,\dots,x_n) = M_1(M_2(\dots(M_{r-1}(B_{r+1}M_rP_c+B_r)+B_{r-1})\dots)+B_2)+B_1,$$

where $M_i = \prod_{i=1}^{k_i} Q_{S_{i_j}}$, $k = k_1 + \cdots + k_r$ is the canalizing depth, P_c is a polynomial that is not k-canalizing, $B_1, B_2, \dots, B_{r+1} \in \mathbb{F}$, and $B_{r+1} \neq 0$. Each variable x_i appears in exactly one of the $M_1, M_2, \dots, M_r, P_c$.

Conclusions

- Algebraic methods are useful for identifying controllers in discrete networks.
- The hierarchy of the canalizing variables can be used for assessing the impact of controllers on the dynamics of the uncontrolled network.
- The upper bound for assessing the impact of the controllers is sharp.
- These two complementary methods can be used for selecting controllers that minimize the side effects resulting from an edge deletion.

Acknowledgments

Thank You!

Collaborators

- Boris Aguilar, Institute for Systems Biology.
- Elena Dimitrova, Clemson University.
- Reinhard Laubenbacher, University of Connecticut Health Center.
- Alan Veliz-Cuba, University of Dayton.

References:

- Identification of control targets in Boolean molecular network models via computational algebra.
  David Murrugarra, Alan Veliz-Cuba, Boris Aguilar, and Reinhard Laubenbacher.
  BMC Systems Biology, 10:94, 2016.
- Molecular Network Control Through Boolean Canalization.
  David Murrugarra and Elena Dimitrova.
  EURASIP Journal on Bioinformatics and Systems Biology, 2015:9, 2015.

email: murrugarra@uky.edu