Bioinformatics 3 V10 – Dynamic **Regulation:** Petri & Boolean Networks

Fri, Nov 22, 2011

Gene Regulation Networks

Biological regulation <=> Projected regulatory network via proteins and metabolites







Dynamic Reconstruction

Different network topologies \rightarrow different time series



Model large networks efficiently → simplified descriptions (processes + numerics)

Quorum sensing of Vibrio fischeri

This luminescent bacterium can be found in small amounts in the ocean and in large amount in isolated areas such as the light organs of squid.

When in small concentrations of cells, *V. fischeri* does not give off light, but in high cell density these bacteria emit a blue-green light. This cell density-dependent control of gene expression is activated by autoinduction that involves the coupling of a transcriptional activator protein with a signal molecule (autoinducer) that is released by the bacteria into its surrounding environment.

In the ocean, the population density of *V. fischeri* is only about 10² cells/ml. The exportation of the autoinducer from the bacteria into this low concentration of cells is not enough to cause the luminescence genes to be activated. However, inside the light organ of a squid for example, the cell concentration is about 10¹⁰ cells/ml. At such high concentrations, the autoinducer causes the bacteria to emit light

> https://www.bio.cmu.edu/courses/03441/TermPapers/99TermPapers/Quorum/vibrio_fischeri.html Bioinformatics III Bioinformatics III

Quorum sensing of Vibrio fischeri

V. fischeri has a microbial symbiotic relationship with the squid Euprymna scolopes. The light organ of the squid provides the bacteria all of the nutrients that they need to survive. The squid benefits from the bacteria's quorum sensing and bioluminescence abilities.

During the day, the squid keeps the bacteria at lower concentrations by expelling some of them into the ocean during regular intervals. At night however, the bacteria are allowed to accumulate to about 10¹⁰ cells/ml so that they will emit blue-green light.

This is perfect for the squid because it is a night feeder. In the moonlight, the swimming squid would normally cast a shadow beneath itself making it a perfect target for squid-eating organisms. However, the bacterial glow will counter the shadowing effect the moon makes and mask the squid from its predators. In the morning, the squid expels some bacteria into the ocean to a concentration where they will not generate light anymore so as to conserve energy_{https://www.bio.cmu.edu/courses/03441/TermPapers/99TermPapers/Quorum/vibrio_fischeri.html}

10. Lecture WS 2013/14

Quorum sensing of Vibrio fischeri



Boolean Networks

"Blackboard explanations" often formulated as conditional transitions

- "If LuxI is present, then AI will be produced..."
- "If there is AI and there's no LuxR:AI bound to the genome, then LuxR will be expressed and complexes can form..."
- "If LuxR:AI is bound to the genome, then LuxI is expressed..."

Simplified mathematical **description** of the dependencies:

- Densities of the species <=> discrete states: on/off, 1/0
- Network of dependencies
- Progress in time

- <=> condition tables
- <=> discrete propagation steps

Boolean Networks II

State of the system: described by vector of discrete values

 $S_i = \{0, 1, 1, 0, 0, 1, \ldots\}$

 $S_i = \{x_1(i), x_2(i), x_3(i), \ldots\}$

fixed number of species with finite number of states each

- \rightarrow finite number of system states
- \rightarrow periodic trajectories

→ periodic sequence of states = attractor

 \rightarrow all states leading to an attractor = **basin of attraction**

Propagation:

$$S_{i+1} = \{x_1(i+1), x_2(i+1), x_3(i+1), \ldots\}$$

$$x_1(i+1) = f_1(x_1(i), x_2(i), x_3(i), \ldots) \quad \text{with } f_i \text{ given}$$

with f_i given by condition tables

A Small Example

State vector $S = \{A, B, C\} \rightarrow 8$ possible states

Conditional evolution:

A is on if C is on $A_{i+1} = C_i$

1

#

0

1

A activates B

 Bi+1
 Ai

 0
 0

 1
 1

С

0

0

1

0

В

0

1

0

0

Α

1

0

Start from {A, B, C} = {1, 0, 0}

Si

 S_0

 S_1

1

C is on if (B is on && A is off)

Ai

0

0

1

Bi

0

1

0

1

B

assume that inhibition through A is stronger than activation via B

periodic orbit of length 3

C_{i+1}

0

1

0

0

Test the Other States



 \rightarrow Either all off or stable oscillations

A Knock-out Mutant





Attractors:





no feedback

 \rightarrow no stabilization, network just "rotates"

Boolean Network of QS



Minimum set of species:

LuxR, AI, LuxR:AI, LuxR:AI:genome, LuxI Here: Light signal (LuxAB) α LuxI

Condition tables: describe the state of a species in the next step given the current states of all relevant species.

Luxl	LuxR:AI:Genome	LuxR:AI:Genome	LuxR:AI
0	0	0	0
1	1	1	1
Ho on	w does LuxI depend LuxR:AI:Genome?	How does LuxR:AI: on LuxR:AI?	Genome depend

10. Lecture WS 2013/14

Condition Tables for QS II

				L	.uxR	LuxR		AI	LuxR:AI:Genome
	A	•			1	0		0	0
LuxR	LuxR				1	1		0	0
		Lux			1	0		1	0
	LuxR				1	1		1	0
	uxR	luxIC	DABE		0	0		0	1
					1	1		0	1
					0	0		1	1
					0	1		1	1
LuxR:AI	LuxR	AI	LuxR:AI:Genome		-				
0	0	0	0		LuxF	R:AI Lu	xR	AI	LuxR:AI:Genome
0	1	0	0					0	
1	0	1	0	\rightarrow	0		X	0	X
1	1	1	0		1		1	1	X
0	0	0	1		1)	1	0
0	1	0	1		0	()	1	1
0	0	1	1		Not	e: no dis	socia	tion	
1	1	1	1		onh	(LuxR	Al:G		$he \rightarrow LuxR:AI + Genome)$
10. Lectu	ure WS 20	13/14	Bioinfo	ormati	cs III	LuxR:	Al:Ge	enom	$e \rightarrow LuxR + Genome$

Condition tables for QS III



AI	LuxR	AI	LuxI
0	0	0	0
0	1	0	0
1	0	1	0
0	1	1	0
1	0	0	1
1	1	0	1
1	0	1	1
1	1	1	1

Α	.i I	_uxR	AI	LuxI
		Х	Х	1
()	Х	0	0
		0	1	0
()	1	1	0

Scanning for Attractors

States of V. fischeri QS system mapped onto integers

{LuxR (LR), LuxR:AI (RA), AI, LuxR:AI:Genome (RAG), LuxI (LI)}

= $\{1, 2, 4, 8, 16\}$ - current state is binary number!

For each attractor:

- periodic orbit and its length (period)
- basin of attraction and its relative size (32 states in total) \rightarrow how likely will the system end in each of the attractors?
- Attractor 1:orbit: 1 \rightarrow period 1states: 0, 1 \rightarrow size 2, 2/32 = 6.25 %



Scanning for Attractors II



averaged occupancies in this periodic orbit:

 LR
 RA
 AI
 RAG
 LI

 4/4 = 1
 1/4 = 0.25
 1/4 = 0.25
 1/4 = 0.25
 1/4 = 0.25

Attractors III

Attractor 3: period 4, basin of 16 states \rightarrow 50 %

LR RA AI RAG LI – state0 . X X . . - 61 . X X X . - 142 . . X X X - 283 . . X . X - 20

Attractor 4: period 4, basin of 4 states \rightarrow 12.5 %

LR RA AI RAG LI - state0 X X X . . - 71 X X . X . . - 112 X . . X X - 253 X . X . X - 21

Attractor 5: period 2, basin of 3 states \rightarrow 9.4 %

LR RA AI RAG LI - state0 X X X - 131 X X - 18

Classifying the Attractors

 \rightarrow Interpret the system's behavior from the properties of the attractors

Attractor	period	basin size	<luxr></luxr>	<luxr:ai></luxr:ai>	<ai></ai>	<luxr:ai:gen></luxr:ai:gen>	<luxl></luxl>
1	1	6.25 % (2)	1	0	0	0	0
2	4	21.9% (7)	1	0.25	0.25	0.25	0.25
3	4	50 % (16)	0	0.5	1	0.5	0.5
4	4	12.5 % (4)	1	0.5	0.5	0.5	0.5
5	2	9.4% (3)	0.5	0.5	0.5	0.5	0.5

Three **regimes**:

dark: LuxI = 0intermediate: LuxI = 0.25bright: LuxI = 0.5free LuxR, no AIfree LuxR + little AIlittle free LuxR (0.24) +
much AI (0.85)

The Feed-Forward-Loop

Y

0

1

0

1

Y

0

1

0

1

External signal determines state of X \rightarrow response Z for short and long signals X



10. Lecture WS 2013/14

Response to signal X(t)

U

Signal propagation

Left column: external signal

Х	Y	Z	
0	0	0	Short
1	0	0	Short
0	1	0	Signal
0	0	0	
1	0	0	
1	1	0	Long
1	1	1	signal
0	1	1	C
0	0	0	
0	0	0	
Х	Y	Z	
0	1	0	
1	1	•	
-	l l	0	
0	0	0	
0	0 1	0 0 0	
0 0 1	0 1 1	0 0 0 0	
0 0 1 1	0 1 1 0	0 0 0 0 0	
0 0 1 1 1	0 1 1 0 0	0 0 0 0 0 1	
0 0 1 1 1 0	0 1 1 0 0 0	0 0 0 0 1 1 1	
0 0 1 1 1 0 0	0 1 1 0 0 0 1	0 0 0 0 1 1 1 1	

U

19

The A. thaliana Flowering Network





Model organism in genomics:

- small, convenient to grow
- completely sequenced (2000): 125 Mbp
- can be easily mutated

also see: Arabidopsis Information Resource (TAIR)@ www.arabidopsis.org/

images from wikimedia

Dynamics of the Genetic Regulatory Network for Arabidopsis thaliana Flower Morphogenesis

LUIS MENDOZA AND ELENA R. ALVAREZ-BUYLLA*

Lab. de Genética Molecular y Evolución, Instituto de Ecología, UNAM, Ap. Postal 70-275, Coyoacán, D.F. CP04510, Mexico

(Received on 26 September 1997, Accepted in revised form on 3 March 1998)

We present a network model and its dynamic analysis for the regulatory relationships among 11 genes that participate in *Arabidopsis thaliana* flower morphogenesis. The topology of the network and the relative strengths of interactions among these genes were based from published genetic and molecular data, mainly relying on mRNA expression patterns under wild type and mutant backgrounds. The network model is made of binary elements and we used a particular dynamic implementation for the network that we call semi-synchronic. Using this method the network reaches six attractors; four of them correspond to observed patterns of gene expression found in the floral organs of *Arabidopsis* (sepals, petals, stamens and carpels) as predicted by the ABC model of flower morphogenesis. The fifth state corresponds to cells that are not competent to flowering, and the sixth attractor predicted by the model is never found in wild-type plants, but it could be induced experimentally. We discuss the biological implications and the potential use of this network modeling approach to integrate functional data of regulatory genes of plant development.

© 1998 Academic Press

J. theor Biol. **193** (1998) 307

10. Lecture WS 2013/14

The ABC Model

Coen, Meyerowitz (1991):

three different activities A, B, and C, active in two adjacent whorls, mutual inhibition of A and C

 \rightarrow combinations determine fate of the tissue



Mendoza, Alvarez-Buylla, J. theor Biol. 193 (1998) 307

ABC Mutants



If any of the 3 functions (activities) is missing, the flowers have different tissue combinations.

> se = sepals, pe = petals, st = stamens, ca = carpels, se* = se, pe, pe

The Network Model

11 genes (including the four ABC genes)



10. Lecture WS 2013/14

Model Implementation

Here: Boolean model with weighted interactions

Propagate state vector $\mathbf{x} = \{x_1, x_2, \dots, x_{11}\}$ by:

$$x_i(t+1) = \mathbf{H}\left(\sum_{j=1}^N w_{ij}x_j(t) - \theta_i\right)$$

Heavyside step function:
$$\mathbf{H}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

Weights w_{ij} and threshold θ_i are not known exactly

- \rightarrow choose integers for simplicity
- \rightarrow positive for activation, negative for inhibition

Mendoza, Alvarez-Buylla, J. theor Biol. 193 (1998) 307

The Numbers

EMF1, TFL1, LFY, AP1, CAL, LUG, UFO, BFU, AG, AP3, PI and SUP.

	-											_		
	0	0	0	0	0	0	0	0	0	0	0	0		0
	1 0 -2 0 0 0	0	0	0	0	0	0	0		0				
	-2	-1	0	2	1	0	0	0	0	0	0	0		3
	-1	0	5	0	0	0	0	0	-1	0	0	0		-1
	0	0	2	0	0	0	0	0	0	0	0	0		1
	0	0	0	0	0	0	0	0	0	0	0	0		0
w =	0	0	0	0	0	0	0	0	0	0	0	0	U =	0
	0	0	0	0	0	0	0	0	0	1	1	0		1
	0	$^{-2}$	1	$^{-2}$	0	-1	0	0	0	0	0	0		-1
	0	0	3	0	0	0	2	1	0	0	0	$^{-2}$		0
	0	0	4	0	0	0	1	1	0	0	0	-1		0
	0	0	0	0	0	0	0	0	0	0	0	0		0

Mendoza, Alvarez-Buylla, J. theor Biol. 193 (1998) 307

Synchronous vs. Asynchronous

Synchronous propagation (Kauffman (1969)):

→ update all species **simultaneously**

 \rightarrow biological problem: do all genes respond at exactly the same time?

Asynchronous propagation (Thomas (1991)):

 \rightarrow update one species after the other in chosen order

 \rightarrow order of update may influence dynamic gene activation patterns

Semi-synchronic propagation (Mendoza (1998)):

 \rightarrow split genes in groups:

 \rightarrow synchronous within group, one group after the other

 \rightarrow base order of groups upon experimental data (it's still a "choice")

 $\textit{EMF1, TFL1} \rightarrow \textit{LFY, AP1, CAL} \rightarrow \textit{LUG, UFO, BFU} \rightarrow \textit{AG, AP3, PI} \rightarrow \textit{SUP}$

Mendoza, Alvarez-Buylla, J. theor Biol. 193 (1998) 307

Some Example Patterns



Exhaustive search: start from all $2^{12} = 4096$ possible initial states,

run for t = 200 steps

 \rightarrow six stationary patterns (attractors of size 1)

Mendoza, Alvarez-Buylla, J. theor Biol. 193 (1998) 307

10. Lecture WS 2013/14

The Attractors



From gene activation patterns in the attractors:

 \rightarrow identify the **four floral** tissue **types** of the ABC model

- \rightarrow one attractor with floral **inhibitors** EMF1, TFL1
 - (characteristic for cells that are not part of the flowers)
- \rightarrow one yet **unidentified** state

Mendoza, Alvarez-Buylla, J. theor Biol. 193 (1998) 307

Possible Pathways



Note: the model does not include temporal and spatial information required to predict where and when which genes are activated or repressed ("signals")

→ these pathways are a "proposal" only

Mendoza et al, Bioinformatics 15 (1999) 593

Sophistication of Networks

A few years later: additional genes and predicted interactions (- - -)



Predictions for Petunia



From *A. thaliana* predict/understand *green petals* mutant phenotype for petunia.

Espinosa-Soto, Padilla-Longoria, Alvarez-Buylla, *The Plant Cell* **16** (2004) 2923

What is it Worth?

Generally: \rightarrow quality of the **results** depends on the quality of the **model**

 \rightarrow quality of the model depends on the quality of the **assumptions**

Assumptions for the Boolean network description:

- (• subset of the species considered
- only discrete density levels
- conditional yes—no causality
- discretized propagation steps

- \rightarrow reduced system state space)
- → dynamic balances lost, reduced to oscillations
 → no continuous processes
- \rightarrow timing of concurrent paths?

"You get what you pay for"

Petri-Nets



Bipartite graph of

- places
- transitions
- directed weighted arcs

two types of nodes



Petri Nets: More Accurate

Places: have a capacity $(1 \dots \infty)$ \rightarrow max. number of tokens (default: ∞)

Arcs: have costs $(1 \dots \infty)$

 \rightarrow number of tokens that are consumed/produced (default: 1)

Transitions: can fire, when the conditions are fulfilled \rightarrow enough tokens on the in-places:

 \rightarrow enough remaining capacity on the out-places:

≥ costs for in-arcs≥ costs for out-arcs



Marking = state of the network = numbers of tokes on the places

Multiple Possibilities



When **multiple transitions** may fire:

- all are equal
- \rightarrow choose one randomly
- if priorities are defined
- \rightarrow transition with highest priority fires



Platform Independent Petri Net Editor



10. Lecture WS 2013/14

Bioinformatics III http://pipe2.sourceforge.net/

"Token Game"



Token Spread

Token Game = stochastic simulation

Run	P1	P2
1	10	10
2	15	5
3	11	9
4	9	11
5	13	7
6	7	13
7	7	13
8	5	15
9	9	11
10	8	12
<n></n>	9.4	10.6
σ	2.8	2.8



for comparison:

expected from Poisson distribution

 $\lambda = 10$ $\sigma = \lambda^{1/2} \approx 3.2$

10. Lecture WS 2013/14

Inhibition



Time Consuming Transitions

Until now: every transition was instantaneous

SPN (Stochastic Petri Net):

Each transition takes some time – exponentially distributed waiting times

$$\frac{dN}{dt} = -kN \implies N(t) = N(0)e^{-kt}$$

=> survival times distributed exponentially



GSPN (Generalized Stochastic Petri Net):

Time-consuming and instantaneous transitions are mixed

DSPN (Deterministic Stochastic Petri Net): Waiting times are fixed or exponentially distributed

General Petri nets: all types of transitions may occur

Petri Nets for Gene Regulation

To encode the dependencies of gene regulation we need: activation, inhibition, logical and, logical or

- transcription factors are not consumed \rightarrow read arcs
- encode on/off states \rightarrow capacity constraints on the places









Boolean Regulatory Petri Nets

Introduce complementary places: tokens on g_1 plus on $\overline{g_1} = 1$ \rightarrow capacity constraints fulfilled automatically (when initial markings are okay) \rightarrow no inhibitory arcs required



Reverse Engineering Networks

Problem: "Find **the** network that explains the biological processes!" \rightarrow usually too ambitious

Experiments: co-expression data

- \rightarrow co-regulation of different genes (correlation or direct interaction?)
- \rightarrow time-series of individual genes

Strategies."Find **all** networks that are compatible with the experiments"

 \rightarrow combinatorial explosion, usually too many candidates

 \rightarrow does not work...

"Find **one** network that is compatible with the experiments"

 \rightarrow solvable task, but how good is this network?

 \rightarrow does not work...

"Find **some** networks that are compatible with the experiments" → algorithms exist, need heuristics (experience) to assess coverage → does work...

Flowering in Arabidopsis



Minimal model of flower morphogenesis in *A. thaliana* \rightarrow only "red" genes

Identify steady states of different parts of the flower

 \rightarrow find **dead** markings

$$\begin{bmatrix} M_d^1(T) = 0 \\ M_d^1(L) = 0 \\ M_d^1(A) = 1 \\ M_d^1(G) = 0 \\ M_d^1(P) = 0 \\ M_d^1(I) = 0 \end{bmatrix} \begin{bmatrix} M_d^2(T) = 0 \\ M_d^2(L) = 0 \\ M_d^2(A) = 1 \\ M_d^2(G) = 0 \\ M_d^2(P) = 1 \\ M_d^2(I) = 1 \end{bmatrix} \begin{bmatrix} M_d^3(T) = 0 \\ M_d^3(L) = 0 \\ M_d^3(A) = 0 \\ M_d^3(G) = 1 \\ M_d^3(P) = 0 \\ M_d^3(I) = 0 \end{bmatrix} \begin{bmatrix} M_d^4(T) = 0 \\ M_d^4(L) = 0 \\ M_d^4(A) = 0 \\ M_d^4(G) = 1 \\ M_d^4(P) = 1 \\ M_d^4(I) = 1 \end{bmatrix}$$

Sepals

Petals

Carpels

Stamens

10. Lecture WS 2013/14

Bioinformatics III Chaouiya et al., *LNCS* **3099** (2004) 137

Drosophila Cell Cycle

Minimal regulation network for the first cell cycles during *D. melanogaster* embryonic development

(MPF = Mitosis Promoting Factor)





Asynchronous graph of all possible states (and transitions) — MFWS

Does the model reproduce **oscillations**?

 \rightarrow prove that the system is **deadlock-free**

(evaluate conditions that any of the transitions cannot fire any more)

Note on a mutation: when MPF inhibits Fizzy \rightarrow dead markings \rightarrow no stable oscillations

Bioinformatics III Chaouiya et al., *LNCS* **3099** (2004) 137

Summary

Today: simplified modelling of gene regulation networks

Boolean Networks

genes are on/off, propagation via condition tables

- \rightarrow direct implementation of experimentally found dependencies
 - \rightarrow no real-time information
 - \rightarrow steady states (attractors) network reconstruction mutations

Petri nets

places, transitions, and arcs (plus capacities)

- \rightarrow more general, more analysis tools, but more complex
 - \rightarrow can include real-time dynamics (via time-consuming transitions)

Next lecture:

graph connectivity