

Modeling and simulation of gene regulatory networks 3

Hidde de Jong IBIS INRIA Grenoble – Rhône-Alpes Hidde.de-Jong@inria.fr

November 14, 2018

INRIA Grenoble - Rhône-Alpes and IBIS



- IBIS: systems biology group at INRIA/Université Joseph Fourier/CNRS
 - Analysis of bacterial regulatory networks by means of models and experiments
 - Biologists, computer scientists, mathematicians, physicists, ...

http://team.inria.fr/ibis

nnia

JNIVERSITÉ



Overview

- 1. Gene regulatory networks in bacteria
- 2. Quantitative modeling of gene regulatory networks
- **3.** Qualitative modeling of gene regulatory networks
- 4. Identification of gene regulatory networks
- 5. Towards integrated models of the cell



Gene regulatory networks

• Gene regulatory networks control changes in gene expression levels in response to environmental perturbations



UNIVERSITÉ Grenoble

nain

Gene regulatory networks consist of genes, gene products, signalling metabolites, and their mutual regulatory interactions

> Global regulators of transcription involved in glucose-acetate diauxie in *E. coli*

Kotte et al. (2010), Mol. Syst. Biol., 6:355

Modeling of gene regulatory networks

• Well-established theory for modeling of gene regulatory networks using ordinary differential equation (ODE) models

Bolouri (2008), *Computational Modeling of Gene Regulatory Networks*, Imperial College Press Polynikis *et al.* (2009), *J. Theor. Biol.*, 261(4):511-30

- Practical problems encountered by modelers:
 - Knowledge on molecular mechanisms rare
 - Quantitative information on kinetic parameters and molecular concentrations absent
 - Large models



Qualitative modeling and simulation

- Intuition: essential properties of network dynamics robust against reasonable model simplifications
- Qualitative modeling and simulation of large and complex gene regulatory networks using simplified models

de Jong, Gouzé et al. (2004), Bull. Math. Biol., 66(2):301-40

• Relation with discrete, logical models of gene regulation Thomas and d'Ari (1990), *Biological Feedback*, CRC Press Kauffman (1993), *The Origins of Order*, Oxford University Press



Ordinary differential equation models

 Gene regulatory networks modeled by ODE models using sigmoid functions to describe regulatory interactions

$$\dot{x}_{a} = \kappa_{a} h^{-}(x_{a}, \theta_{a2}, n) h^{-}(x_{b}, \theta_{b}, n) - \gamma_{a} x_{a}$$
$$\dot{x}_{b} = \kappa_{b} h^{-}(x_{a}, \theta_{a1}, n) - \gamma_{b} x_{b}$$





- x : protein concentration
- θ : threshold concentration
- κ , γ : rate constants
- n : steepness parameter
- Expressions of sigmoid functions account for combinatorial control of gene expression (AND, OR, NOR, ...)



PL differential equation models

 ODE models approximated by means of step functions to describe regulatory interactions

$$\dot{x}_{a} = \kappa_{a} \, s^{-}(x_{a}, \theta_{a2}) \, s^{-}(x_{b}, \theta_{b}) - \gamma_{a} \, x_{a}$$
$$\dot{x}_{b} = \kappa_{b} \, s^{-}(x_{a}, \theta_{a1}) - \gamma_{b} \, x_{b}$$





- x: protein concentration
- θ : threshold concentration
- κ , γ : rate constants
- Piecewise-linear (PL)DE models of gene regulatory networks

Glass and Kauffman (1973), J. Theor. Biol., 39(1):103-29



• Analysis of local dynamics of PL models

Monotone convergence towards **focal point** in regions separated by thresholds



Glass and Kauffman (1973), J. Theor. Biol., 39(1):103-29



• Analysis of local dynamics of PL models

Monotone convergence towards **focal point** in regions separated by thresholds



Glass and Kauffman (1973), J. Theor. Biol., 39(1):103-29



Analysis of local dynamics of PL models
 Instantaneous crossing of regions located on thresholds, or ...



$$\dot{x}_{a} = \kappa_{a} s^{-}(x_{a}, \theta_{a2}) s^{-}(x_{b}, \theta_{b}) - \gamma_{a} x_{a}$$
$$\dot{x}_{b} = \kappa_{b} s^{-}(x_{a}, \theta_{a1}) - \gamma_{b} x_{b}$$



- Analysis of local dynamics of PL models
 - ... quasi-monotone convergence towards **focal sets** located on threshold hyperplanes max_h



$$\dot{x}_b = \kappa_b \, s(x_a, \theta_{a1}) - \gamma_b \, x_b$$

 Extension of PL differential **equations** to differential **inclusions** using Filippov approach Gouzé and Sari (2002), *Dyn. Syst.*, 17(4):299-316



• Analysis of global dynamics obtained by piecing together local dynamics in regions

PL approximation preserves bistability of cross-inhibition network





- State space can be partitioned into regions with unique derivative sign pattern
- **Qualitative abstraction** yields state transition graph that provides discrete picture of continuous dynamics

Alur et al. (2000), Proc. IEEE, 88(7):971-84



de Jong *et al.* (2004), *Bull. Math. Biol.*, 66(2):301-40 Batt *et al.* (2008), *Automatica*, 44(4):982-9



- State transition graph gives **conservative approximation** of continuous dynamics
 - Every solution of PL model corresponds to path in state transition graph
 - Converse is not necessarily true!
- State transition graph is **invariant** for given inequality constraints on parameters



- State transition graph gives conservative approximation of continuous dynamics
 - Every solution of PL model corresponds to path in state transition graph
 - Converse is not necessarily true!
- State transition graph is **invariant** for given inequality constraints on parameters



- State transition graph gives **conservative approximation** of continuous dynamics
 - Every solution of PL model corresponds to path in state transition graph
 - Converse is not necessarily true!
- State transition graph is **invariant** for given inequality constraints on parameters



$$0 < \kappa_{a}/\gamma_{a} < \theta_{a1} < \theta_{a2} < max_{a}$$
$$0 < \theta_{b} < \kappa_{b}/\gamma_{b} < max_{b}$$

Batt et al. (2008), Automatica, 44(4):982-9

Use of state transition graph

- Analysis of steady states and limit cycles of PL models
 - Attractor states in graph correspond (under certain conditions) to stable steady states of PL model
 Casey et al. (2006), J. Math Biol., 52(1):27-56
 - Attractor cycles in graph correspond (under certain conditions) to stable limit cycles of PL model

Glass and Pasternack (1978), *J. Math Biol.*, 6(2):207-23 Edwards (2000), *Physica D*, 146(1-4):165-99



Use of state transition graph

- Paths in state transition graph represent predicted sequences of qualitative events
- Model validation: comparison of predicted and observed sequences of qualitative events



Need for automated and efficient tools for model validation

Model validation by model checking

Dynamic properties of system can be expressed in temporal logic (CTL)

There Exists a Future state where $\dot{x}_a > 0$ and $\dot{x}_b > 0$ and starting from that state, there Exists a Future state where $\dot{x}_a < 0$ and $\dot{x}_b > 0$

$$\mathrm{EF}(\dot{x}_a > 0 \land \dot{x}_b > 0 \land \mathrm{EF}(\dot{x}_a < 0 \land \dot{x}_b > 0))$$



 Model checking is automated technique for verifying that state transition graph satisfies temporal-logic statements
 Efficient computer tools available for model checking

Batt et al. (2005), Bioinformatics, 21(supp. 1): i19-i28



Genetic Network Analyzer (GNA)

Qualitative analysis of PL models implemented in Java: Genetic
 Network Analyzer (GNA)



http://www-helix.inrialpes.fr/gna



Genetic Network Analyzer (GNA)

- Model-checking technology made available to GNA user
 - Develop temporal logics tailored to biological questions

Mateescu *et al.* (2011), *Theor. Comput. Sci.*, 412:2854-83

 Develop temporal-logic patterns for frequentlyasked modeling questions



Monteiro et al. (2008), Bioinformatics, 24(16):i227-33

 Connect GNA to standard model checkers through a web-server connection

Monteiro et al., (2009), BMC Bioinform., 10:450



Analysis of bacterial regulatory networks

- Applications of qualitative simulation in bacteria:
 - Initiation of sporulation in Bacillus subtilis

de Jong, Geiselmann et al. (2004), Bull. Math. Biol., 66(2):261-300

 Quorum sensing in *Pseudomonas* aeruginosa

> Viretta and Fussenegger (2004), *Biotechnol. Prog.*, 20(3):670-8

 Onset of virulence in *Erwinia* chrysanthemi

Sepulchre et al. (2007), J. Theor. Biol., 244(2):239-57



Lasl



Biodegradation of polluants by P. putida

 Soil bacterium *Pseudomonas putida* mt-2 is archetypal model for environmental biodegradation of aromatic pollutants

TOL network involved in degradation of *m*-xylene to intermediates for central carbon metabolism



Rocha-Silva et al. (2011), Environ. Microbiol., 13(9):2389-402



Role of regulators of TOL network

• **Question**: what is the role of the central, plasmid-encoded regulators XyIR and XyIS?



Development of PL model of TOL network

Translation of network diagram into regulatory logic and PL model

Rocha-Silva et al. (2011), BMC Syst. Biol., 5:191



Role of regulators of TOL network

• Validation of model by testing predictions under different perturbation conditions (mutants, metabolic inducers, ...)



 Plasmid-encoded regulators of TOL network act as regulatory firewall

Prevent toxic *m*-xylene and its biodegradation intermediates from intervening with indigenous metabolic pathways

Rocha-Silva et al. (2011), BMC Syst. Biol., 5:191



IRMA: synthetic network in yeast

• IRMA: synthetic network in yeast consisting of interlocked positive and negative feedback loops

Networks functions independently from host cell

 Network can be externally controlled by growing cells in glucose or galactose



Cantone et al. (2009), Cell, 137(1):172-81



IRMA: synthetic network in yeast

- IRMA proposed as a benchmark for modeling and identification approaches
- IRMA dynamics measured over time in galactose (switch-on) and glucose (switch-off) Quantitative RT-PCR
- Question: are measured dynamics consistent with constructed network structure?



Cantone et al. (2009), Cell, 137(1):172-81



- Development of (unparametrized) PL model representing network structure
- Approach to test consistency between network structure and data based on automated parameter constraint search:
 - Generate temporal logic formulae encoding observed network dynamics



Batt *et al.* (2010), *Bioinformatics*, 26(18):i603-10



- Development of (unparametrized) PL model representing network structure
- Approach to test consistency between network structure and data based on automated parameter constraint search:
 - Generate temporal logic formulae encoding observed network dynamics
 - Test if there are any parametrizations of PL model satisfy temporal logic formulae

		Symbolic state space and symbolic parameter space		Symbolic state space and explicit parameter space	
Г	Property	Existence of	Parametrization*	Number of	Parametrization*
L		parametrization		parametrizations	
	ϕ_1 : averaged time-series	Yes (49 s)	$\frac{\kappa_{Swi5}^{0}}{\gamma_{Swi5}} < \theta_{Swi5}^{g} < \theta_{Swi5}^{c} < \theta_{Swi5}^{a} < \frac{\kappa_{Swi5}^{0} + \kappa_{Swi5}}{\gamma_{Swi5}} \\ \wedge \frac{\kappa_{Gal80}^{0}}{\gamma_{Gal80}} < \frac{\kappa_{Gal80}^{0} + \kappa_{Gal80}}{\gamma_{Gal80}} < \theta_{Gal80}$	12 (925 s)	$ \begin{array}{c} \frac{\kappa_{Swi5}^{0}}{\gamma_{Swi5}} < \theta_{Swi5}^{c} < \theta_{Swi5}^{a} < \frac{\kappa_{Swi5}^{0} + \kappa_{Swi5}}{\gamma_{Swi5}} \land \\ (\theta_{Gal80} < \frac{\kappa_{Gal80}^{0}}{\gamma_{Gal80}} \land \frac{\kappa_{Swi5}^{0}}{\gamma_{Swi5}} < \theta_{Swi5}^{g} < \frac{\kappa_{Swi5}^{0} + \kappa_{Swi5}}{\gamma_{Swi5}} \\ \lor \frac{\kappa_{Gal80}^{0}}{\gamma_{Gal80}^{c}} < \theta_{Gal80} < \frac{\kappa_{Gal80}^{0} + \kappa_{Gal80}}{\gamma_{Gal80}} \land \frac{\kappa_{Swi5}^{0}}{\gamma_{Swi5}} < \theta_{Swi5}^{g} \\ \lor \frac{\kappa_{Gal80}^{0} + \kappa_{Gal80}}{\gamma_{Gal80}^{c}} < \theta_{Gal80} \\ < \frac{\kappa_{Gal80}^{0}}{\gamma_{Gal80}^{c}} < \theta_{Gal80} \\ \end{array} \right) $

*All parametrizations additionally include $\kappa_{Cbf1}^1/\gamma_{Cbf1} < \theta_{Cbf1} < (\kappa_{Cbf1}^1 + \kappa_{Cbf1}^2)/\gamma_{Cbf1} \wedge \kappa_{Gal4}^0/\gamma_{Gal4} < \theta_{Gal4} < (\kappa_{Gal4}^0 + \kappa_{Gal4})/\gamma_{Gal4} \wedge \kappa_{Ash1}^0/\gamma_{Ash1} < \theta_{Ash1} < (\kappa_{Ash1}^0 + \kappa_{Ash1})/\gamma_{Ash1}.$



- Development of (unparametrized) PL model representing network structure
- Approach to test consistency between network structure and data based on automated parameter constraint search:
 - Generate temporal logic formulae encoding observed network dynamics
 - Test if there are any parametrizations of PL model satisfy temporal logic formulae
 - Analyze parametrizations for biological plausibility

« Activation threshold of CBF1 by Swi5 higher than activation threshold of ASH1 »: confirmed by independent experimental data

Batt *et al.* (2010), *Bioinformatics*, 26(18):i603-10



- Development of (unparametrized) PL model representing network structure
- Approach to test consistency between network structure and data based on automated parameter constraint search:
 - Generate temporal logic formulae encoding observed network dynamics
 - Test if there are any parametrizations of PL model satisfy temporal logic formulae
 - Analyze parametrizations for biological plausibility
- Automated approach for testing consistency based on modelchecking techniques

Symbolic encoding of model, dynamics and properties to make problem feasible



PL differential equation models

 ODE models approximated by means of step functions to describe regulatory interactions

$$\dot{x}_{a} = \kappa_{a} \, s^{-}(x_{a}, \theta_{a2}) \, s^{-}(x_{b}, \theta_{b}) - \gamma_{a} \, x_{a}$$
$$\dot{x}_{b} = \kappa_{b} \, s^{-}(x_{a}, \theta_{a1}) - \gamma_{b} \, x_{b}$$





- x: protein concentration
- θ : threshold concentration
- κ , γ : rate constants
- Expressions of step functions account for combinatorial control of gene expression (AND, OR, NOR, ...)



PL differential equation models

 ODE models approximated by means of step functions to describe regulatory interactions

$$\dot{x}_a = \kappa_a \, s^{-}(x_b, \theta_b) - \gamma_a \, x_a$$
$$\dot{x}_b = \kappa_b \, s^{-}(x_a, \theta_a) - \gamma_b \, x_b$$





- x : protein concentration
- θ : threshold concentration
- κ , γ : rate constants



Boolean models

 Boolean models are discrete models of dynamics of gene regulatory networks

Thomas and d'Ari (1990), *Biological Feedback*, CRC Press Kauffman (1993), *The Origins of Order*, Oxford University Press Wang et al. (2012), *Phys. Biol.*, 9(5):055001

Boolean variables discretize state of gene regulatory network

 $X_{a} \in \{0, 1\}, X_{b} \in \{0, 1\}$ $X_{a} = (x_{a} > \theta_{a}), X_{b} = (x_{b} > \theta_{b})$ $X_{a}^{t}, X_{b}^{t}, t = 0, 1, 2, \dots$





Boolean models

Boolean models are discrete models of dynamics of gene regulatory networks

Thomas and d'Ari (1990), *Biological Feedback*, CRC Press Kauffman (1993), *The Origins of Order*, Oxford University Press Wang et al. (2012), *Phys. Biol.*, 9(5):055001

- Boolean variables discretize state of gene regulatory network
- Boolean functions represent control of gene expression

 $X_a^{t+1} = \text{NOT } X_b^t$ $X_b^{t+1} = \text{NOT } X_a^t$

nnía



Analysis of Boolean models

• Boolean models can be analyzed in discrete state space



$$X_a^{t+1} = \text{NOT } X_b^t$$
$$X_b^{t+1} = \text{NOT } X_a^t$$

• Synchrone and asynchrone dynamics





Analysis of Boolean models

- Dynamics of Boolean models can also be represented in state transition graph
 - Different graphs for synchrone and asynchrone dynamics
 - Attractors (states or cycles)



nnia



 Generalized logical models are discrete models of dynamics of gene regulatory networks

Thomas and d'Ari (1990), *Biological Feedback*, CRC Press Chaouiya *et al.* (2012), *Methods Mol. Biol.*, 804:463-79

Logical variables discretize state of gene regulatory network

$$X_{a} \in \{0, 1, 2, ...\}, X_{b} \in \{0, 1, 2, ...\}$$

$$X_{a}^{t}, X_{b}^{t}, t = 0, 1, 2, ...$$

$$1$$

$$gene a$$

$$gene b$$

$$0$$

$$1$$

$$0$$

$$1$$

$$2$$

 Generalized logical models are discrete models of dynamics of gene regulatory networks

Thomas and d'Ari (1990), *Biological Feedback*, CRC Press Chaouiya *et al.* (2012), *Methods Mol. Biol.*, 804:463-79

- Logical variables discretize state of gene regulatory network
- Boolean functions represent control of gene expression

$$X_{a}^{t+l} = 2, \text{ if } (X_{a}^{t} = 0 \text{ OR } X_{a}^{t} = 1) \text{ AND } X_{b}^{t} = 0$$

$$X_{a}^{t+l} = 0, \text{ if } X_{a}^{t} = 2 \text{ OR } X_{b}^{t} = 1$$

$$X_{b}^{t+l} = 1, \text{ if } X_{a}^{t} = 0$$

$$X_{b}^{t+l} = 0, \text{ if } X_{a}^{t} = 1 \text{ OR } X_{a}^{t} = 2$$

$$0$$



2

0

 Generalized logical models are discrete models of dynamics of gene regulatory networks

Thomas and d'Ari (1990), *Biological Feedback*, CRC Press Chaouiya *et al.* (2012), *Methods Mol. Biol.*, 804:463-79

- Logical variables discretize state of gene regulatory network
- Boolean functions represent control of gene expression
- Dynamics can be represented by state transition graph Attractors (states and cycles)



asynchrone



 Generalized logical models are discrete models of dynamics of gene regulatory networks

Thomas and d'Ari (1990), *Biological Feedback*, CRC Press Chaouiya *et al.* (2012), *Methods Mol. Biol.*, 804:463-79

- Logical variables discretize state of gene regulatory network
- Boolean functions represent control of gene expression
- Dynamics can be represented by state transition graph
- Close correspondence between discrete abstractions of PLDE models and generalized logical models



GinSIM

GinSIM: computer tool for logical modeling of regulatory networks



Chaouiya et al. (2012), Methods Mol. Biol., 804:463-79



• Development of *Drosophila melanogaster* (fruit fly)



Purves et al. (1998), Life: The Science of Biology, Sinauer





• Development of *Drosophila melanogaster* (fruit fly)



Purves et al. (1998), Life: The Science of Biology, Sinauer

Tomer et al. (2012), Nat. Methods, 9(7):755-63





Grenoble

Alpes

Spatiotemporal gene expression patterns during early development of Drosophila (fruit fly)

Sanson (2001), EMBO Rep., 2(12):1083-8



- Spatiotemporal gene expression patterns during early development of *Drosophila* (fruit fly)
- Gene classes and their interactions responsible for establishment of gene expression patterns

Schroeder et al. (2004), PLoS Biol., 4(2):e271



Carroll (2008), Cell, 134(1):25-36

- Spatiotemporal gene expression patterns during early development of *Drosophila* (fruit fly)
- Gene classes and their interactions responsible for establishment of gene expression patterns
- Complex gene regulatory networks



Model of Drosophila segmentation

 Model of network of segment polarity genes in early development of *Drosophila*





von Dassow et al. (2000), Nature, 406(6792): 188-92



Model of Drosophila segmentation

 Model of network of segment polarity genes in early development of *Drosophila*

13 ODEs per cell and 48 parameters





von Dassow et al. (2000), Nature, 406(6792): 188-92



Robustness of gene expression patterns

 Spatial expression pattern of segment polarity genes robustly reproduced over large ranges of parameter values

0.5% of sampled parameter combinations leads to solution compatible with data



von Dassow et al. (2000), Nature, 406(6792): 188-92



Logical model of Drosophila segmentation

• Logical model of segment polarity network: variables take values 0/1 and Boolean functions to update variables



UNIVERSITÉ

Albert and Othmer (2003), J. Theor. Biol., 223(1):1-18

hh_i HH_i ptc_i PTC_i PH_i

nnia



Logical model of Drosophila segmentation

- Logical model of segment polarity network: variables take values 0/1 and Boolean functions to update variables
- Logical models are based on topology of network only (no parametrization), but are capable of reproducing experimental data: robustness





Conclusions

- Modeling of genetic regulatory networks in bacteria often hampered by lack of information on parameter values
- Use of coarse-grained discrete or discretized models that provide reasonable approximation of dynamics
- Mathematical methods and computer tools for analysis of qualitative dynamics of discrete models
- Use of discrete models may gain insight into functioning of large and complex networks
- Discrete, coarse-grained models provide first idea of qualitative dynamics that may guide quantitative modeling



Merci!



team.inria.fr/ibis