Méthodes de modélisation informatique des processus cellulaires: cycle cellulaire et horloge circadienne

François Fages Institut National de Recherche en Informatique et Automatique Inria Saclay – Ile de France **EP LIFEWARE** http://lifeware.inria.fr/ **BIOCHAM modeling and analysis software** http://lifeware.inria.fr/biocham4

From Biological Facts...

Biology is still mainly a collection of facts

Large amount of data but still few principles and processes really understood:

- Darwin's principle of evolution and natural selection in 1859
- Jakob and Monod's elucidation of transcription and translation in 1961
- Multistability in gene networks explaining cell differentiation, cell reprogramming
- **Oscillations** in gene or protein networks explaining cellular circadian clock, cell cycle, homeostasis, morphogenesis,...
- **Signaling** networks (e.g. MAPK) performing Analog-Digital signal conversion, noise filtering, ultrasensitivity and signal amplification

Need of formal methods to master the complexity of large interaction networks...



From Computer Science Methods...

Alan Turing created Computer Science in 1936 in the perspective of artificial intelligence:

« The behavior of a human doing calculations is at each instant determined by the symbols he observes and by his internal mental state »

Church-Turing Thesis states that there is **one single notion of computation** and one notion of **universal machine**. Indeed

- Turing machines, Church's lambda calculus, Gödel's first-order logic
- Random Access Memory machines
- Fortran, C, Java, CAML, Python etc. programming languages all have the same computational power.

According to this Thesis, any form of computation is thus necessarily representable by a program



... to Cell Processes as Computation

The behavior of a living cell is at each instant determined by the signals it observes and by its internal state.

By applying Church-Turing Thesis to biology, **cell compute**; In the computational metaphor, a cell, a tissue, continuously compute.

The paradigm of « **Cells as Machines** » opens up the methods of Computer Science developped in the last decades for mastering the complexity of

- Circuits of trillions of transistors, programs of millions of instructions,
- World wide web of tenths of millions of web sites

to tackle the complexity of

- tenths of thousands of genes and proteins in a cell
- millions of cells.



Overview of the Course

1. Rule-based modeling of biochemical reaction systems

- Syntax: Biocham notations (SBML compatible)
- Semantics: Differential, Stochastic and Boolean interpretations of reactions
- Examples of cell cycle control, circadian clock, gene expression

2. Temporal Logic based formalization of biological properties

- Quantitative properties in Linear Time Logic LTL(R)
- Parameter search in high dimension w.r.t. LTL(R) specifications
- Robustness and sensitivity analyses w.r.t. LTL(R) specifications

3. Coupled model of the cell cycle and circadian clock

- chronotherapy optimization



Computational Systems Biology

"Systems Biology aims at systems-level understanding which requires a set of principles and methodologies that links the behaviors of molecules to systems characteristics and functions."

H. Kitano, ICSB 2000

- Follow-up of the human genome project of the 90's
- Analyze post-genomic data produced with high-throughput technologies and made available in public databases like GO, KEGG, BioCyc, etc.;
- Integrate heterogeneous data about a specific problem;
- Predict the behavior of large networks of genes and proteins;
- Multi-scale models of cell processes, tissues, organisms, ecosystems...
- → Systems Biology Markup Language (SBML): model exchange format
- → SBML model repositories: e.g. biomodels.net thousands of models



Biochemical reaction models

Models for representing knowledge : the more detailed the better
 Models for answering questions : the more abstract the better



Mammalian cell cycle control interaction map [Kohn 1999]

Simplified yeast cell cycle model [Tyson 1991]

- Binding, complexation: $A + B \rightarrow C$ $cdk1 + cycB \rightarrow cdk1cycB$
- Unbinding, decomplexation:

+

 $A \rightarrow B + C$





 \leftrightarrow





- Binding, complexation: $A + B \rightarrow C$ $cdk1 + cycB \rightarrow cdk1cycB$
- Unbinding, decomplexation: $A \rightarrow B + C$
- Transformation, phosphorylation, transport: $A \rightarrow B$ $(A + K \rightarrow C \rightarrow B + K)$ $cdk1cycB \rightarrow cdk1cycBp$





- Binding, complexation: $A + B \rightarrow C$ $cdk1 + cycB \rightarrow cdk1cycB$
- Unbinding, decomplexation: $A \rightarrow B + C$
- Transformation, phosphorylation, transport: $A \rightarrow B$ $(A + K \rightarrow C \rightarrow B + K)$ $cdk1cycB \rightarrow cdk1cycBp$
- Gene expression, synthesis: $A \rightarrow$

 $A \rightarrow A + B$

NA polymerase DNA binding domain MPCE 2018 promoter site

 $E2Fa \rightarrow E2Fa + RNAcycA$



- Binding, complexation: $A + B \rightarrow C$ $cdk1 + cycB \rightarrow cdk1cycB$
- Unbinding, decomplexation: $A \rightarrow B + C$
- Transformation, phosphorylation, transport: $A \rightarrow B$ $(A + K \rightarrow C \rightarrow B + K)$ $cdk1cycB \rightarrow cdk1cycBp$
- Gene expression, synthesis: $A \rightarrow A + B$ $E2Fa \rightarrow E2Fa + RNAcycA$
- Degradation:



Biochemical reaction ratesTime mattersBinding, complexation: $A + B \xrightarrow{k.A.B} - C$

• Unbinding, decomplexation:

 $cdk1 + cycB \rightarrow cdk1cycB$

- Transformation, phosphorylation, transport:
 cdk1cycB → cdk1cycBp
- Gene expression, synthesis: $E2Fa \rightarrow E2Fa + RNAcycA$
- Degradation:

$$A \xrightarrow{\nu.A/(k+A)} B$$

 $A \xrightarrow{k.A} B + C$

$$A \xrightarrow{\nu \cdot A^n / (k + A^n)} A + B$$

 $A \xrightarrow{k.A}$

MPCE 2018

- Unbinding, decomplexation: **Transformation, phosphorylation, transport**: $cdk1cycB \rightarrow cdk1cycBp$
- Gene expression, synthesis: $E2Fa \rightarrow E2Fa + RNAcycA$

Binding, complexation:

 $cdk1 + cycB \rightarrow cdk1cycB$

• Degradation:

Biochemical reaction rates Time matters $A + B \xrightarrow{k.A.B} C$ dk1cycBmplexation: $A \xrightarrow{k.A} B + C$

 $A \xrightarrow{v.A/(k+A)} B$

 $A \xrightarrow{v.A^n/(k+A^n)} A + B$

 $A \xrightarrow{k.A}$

Biochemical reaction rates Time matters $A + B \xrightarrow{k.A.B} C$ Binding, complexation: Mass action law kinetics $cdk1 + cycB \rightarrow cdk1cycB$ $A \xrightarrow{k.A} B + C$ Michaelis Menten kinetics Unbinding, decomplexation: $A \xrightarrow{v.A/(k+A)} B \qquad (A+K \to C \to B+K)$ Transformation, phosphorylation, transport: $cdk1cvcB \rightarrow cdk1cvcBp$

- Gene expression, synthesis: $E2Fa \rightarrow E2Fa + RNAcycA$
- Degradation:



 $A \xrightarrow{v.A^n/(k+A^n)} A + B$

 $A \xrightarrow{k.A}$



Degradation:



Semantics of Reaction Programs $A + B \xrightarrow{k.A.B} C$

Continuous semantics: concentrations, continuous time evolution

Ordinary differential equations (ODE)

$$\frac{dA}{dt} = -k.A.B \quad \frac{dB}{dt} = -k.A.B \quad \frac{dC}{dt} = k.A.B$$





Semantics of Reaction Programs $A + B \xrightarrow{k.A.B} C$

Continuous semantics: concentrations, continuous time evolution

Ordinary differential equations (ODE)

$$\frac{dA}{dt} = -k.A.B \quad \frac{dB}{dt} = -k.A.B \quad \frac{dC}{dt} = k.A.B$$

Stochastic semantics: numbers of molecules, probability and time of transition

Continuous Time Markov Chain (CTMC)

A, B
$$\xrightarrow{p(S_i), t(S_i)}$$
 C++, A--, B--

 (\mathbf{C})



k.A.B Semantics of Reaction Programs $A + B \longrightarrow C$

Continuous semantics: concentrations, continuous time evolution

 $\frac{dA}{dt} = -k.A.B \quad \frac{dB}{dt} = -k.A.B \quad \frac{dC}{dt} = k.A.B$ Ordinary differential equations (ODE)

Stochastic semantics: numbers of molecules, probability and time of transition A. $B \xrightarrow{p(S_i), t(Si)} C++, A--, B--$ Continuous Time Markov Chain (CTMC)

Multi-agent simulation: numbers of molecules, space, diffusion speed, affinity







Semantics of Reaction Programs $A + B \xrightarrow{k.A.B} C$

Continuous semantics: concentrations, continuous time evolution

Ordinary differential equations (ODE)

 $\frac{dA}{dt} = -k.A.B \quad \frac{dB}{dt} = -k.A.B \quad \frac{dC}{dt} = k.A.B$

 $A : B \xrightarrow{p(S_i), t(Si)} C++, A--, B--$

Stochastic semantics: numbers of molecules, probability and time of transition

Continuous Time Markov Chain (CTMC)

Petri net semantics: numbers of molecules Multiset rewriting CHAM [Berry Boudol 90] [Banatre Le Metayer 86] P-systems [Paun 98]

A, $B \rightarrow C++$, A--, B--



Semantics of Reaction Programs $A + B \xrightarrow{k.A.B} C$

Continuous semantics: concentrations, continuous time evolution

Ordinary differential equations (ODE) $\frac{dA}{dt} = -k.A.B \quad \frac{dB}{dt} = -k.A.B \quad \frac{dC}{dt} = k.A.B$

Stochastic semantics: numbers of molecules, probability and time of transition

Continuous Time Markov Chain (CTMC)

Boolean semantics: presence/absence Asynchronous transition system



A, $B \rightarrow C++$, A--, B--

 $A : B \xrightarrow{p(S_i), t(Si)} C++, A--, B--$

 $A \land B \rightarrow C \land \neg A \land \neg B$ $A \land B \rightarrow C \land A \land \neg B$ $A \land B \rightarrow C \land \neg A \land B$ $A \land B \rightarrow C \land \neg A \land B$

Hierarchy of Semantics: Abstractions



MPCE 2018

Hierarchy of Semantics: Approximations



Informatics mathematics

Minimal Mitotic Oscillator [Goldbeter 91 PNAS]

Cyclin is synthesized at constant rate

Cyclin triggers the activation of Cdc2 kinase M (by dephosphorylation)

Cdc2 activates a protease X (by phosphorylation) that degrades the Cyclin





Demo Biocham Notebook

http://lifeware.inria.fr:8888/notebooks/examples/

Open MPCE file Load oscillator.ipynb



Single Enzymatic Reaction

An enzyme E binds to a substrate S to catalyze the formation of product P:

 $E+S \rightarrow^{k_1} C \rightarrow^{k_2} E+P$ $E+S \leftarrow^{km_1} C$

Mass action law kinetics ODE:

$$\label{eq:dE/dt} \begin{split} dE/dt &= -k_1ES + (k_2 + k_{m1})C \\ dS/dt &= -k_1ES + k_{m1}C \\ dC/dt &= k_1ES - (k_2 + k_{m1})C \\ dP/dt &= k_2C \end{split}$$
 with two conservation laws: E+C=constant, S+C+P=constant, Assuming C_0=P_0=0, we get E=E_0-C and S_0=S+C+P \\ dS/dt &= -k_1(E_0-C)S + k_{m1}C \\ dC/dt &= k_1(E_0-C)S - (k_2 + k_{m1})C \end{split}



Quasi-Steady State Approximation

When E<<S, $k_1 >> k_2$ we have dC/dt~0~dE/dt

 $C = E_0 S/(K_m + S)$ where $K_m = (k_2 + k_{m1})/k_1$

 $dP/dt = -dS/dt = V_mS / (K_m+S)$ where $V_m = k_2E_0$



Leonor Michaelis Maud Menten 1913

<u>Michaelis-Menten kinetics</u>: V_mS /(K_m+S) for S => P

Model reduction: the variables E and C are eliminated, E is supposed constant and acting on only one substrate



Quasi-Steady State Approximation

When E<<S, $k_1 >> k_2$ we have dC/dt~0~dE/dt

 $C = E_0 S/(K_m + S)$ where $K_m = (k_2 + k_{m1})/k_1$ K_m substrate concentration with half maximum velocity

 $dP/dt = -dS/dt = V_mS / (K_m+S)$ where $V_m = k_2E_0$ Let V_m maximum velocity at saturing substrate concentration

<u>Michaelis-Menten kinetics</u>: $V_m S / (K_m + S)$ for S => P

Model reduction: the variables E and C are eliminated, E is supposed constant and acting on only one substrate



Leonor Michaelis Maud Menten 1913



Substrate concentration [S]



Cell Division Cycle



Cell Division Cycle Control





Mammalian Cell Cycle Control Map [Kohn 99]





Cell Cycle Model [Qu McMillan Weiss 03]



Demo Biocham Notebook

http://lifeware.inria.fr:8888/notebooks/examples/

Open MPCE file Load Qu.ipynb



Control of the Cell Cycle by the Circadian CLock

- Time gating for mitosis by effects of clock genes on cell cycle genes inhibition of Wee1 synthesis by Clock-Bmal1 [Matsuo et al 2003]
- Model-based predictions on conditions of entrainment [Calzone Soliman 2006] and period doubling (24h, 48h) phenomena [Gerard Goldbeter 2012] (also repression of c-Myc by Clock-Bmal1 and inhibition of p21 by Reverb-α)



Circadian Clock Model [Leloup Goldbeter 03]





Demo Biocham Notebook

http://lifeware.inria.fr:8888/notebooks/examples/

Open MPCE file Load clock.ipynb



Coupled Model: time gating for mitosis



Coupling through Wee1 [Matsuo et al 2003] [Calzone Soliman 2006]

=> mWee1 at rate (ksweemp+ksweem*[Bmal1N])/(Kweem+kwpcn*[PerCryN])

MPCE 2018

Conditions of Entrainment



- Conditions of entrainment on Bmal1-Wee1 and MPF activation parameters
- Period doubling (24h, 48h) phenomena

[Gerard Goldbeter PLOS 2012]



Irinotecan Exposure Chronotherapy Model



Whole body PK/PD drug injection model [Ballesta et al PlosCB 2011]

MPCE 2018



Specification of Temporal Behaviors

- Linear Time Temporal logic (LTL) extends classical logic with time operators
 X: next, F: finally, G: globally, U: until
- FO-LTL(R_{lin}), can express quantitative properties :
 - Stability **G**φ
 - Reachability $F\phi$, thresholds F([A]>0.1),
 - Peaks of concentration F([A]<V \land X([A]=V \land X([A]<V)) \uparrow
 - Amplitude
 - Periods and phases as distance between peaks
 - ..
- More flexible than curve fitting, can abstract from uncertain imprecise data
- Verification of reachability constraints w.r.t. qualitative observations
- Constraints for parameter search w.r.t. quantitative observations
 - Good fit \rightarrow model-based predictions, control
 - No good fit \rightarrow revisit structure \rightarrow model-based contribution to biology







Naïve Parameter Scanning Algorithm

Input: an ODE model M(p) with n parameters p in range [pmin,pmax], an LTL(R) specification φ
 Output: parameter values v such that M(v) |= φ or fail if no such values

- 1. Scan the parameter value space [pmin,pmax]ⁿ with a fixed step
- 2. Test whether $M(v) \models \phi$ by trace-based model checking
- 3. Return the first value set v which satisfies ϕ

Exponential complexity in O(sⁿ) where n is the number of parameters and s is the number of discrete values to try for each parameter



biocham: search_parameters([k3,k4],[(0,200),(0,200)],20, oscil(Cdc2-Cyclin~{p1},3),150).

First values found : parameter(k3,10). parameter(k4,70).





biocham: search_parameters([k3,k4],[(0,200),(0,200)],20, oscil(Cdc2-Cyclin~{p1},3) & F([Cdc2-Cyclin~{p1}]>0.15), 150).

First values found : parameter(k3,10). parameter(k4,120).





biocham: search_parameters([k3,k4],[(0,200),(0,200)],20, oscil(Cdc2-Cyclin~{p1},3) & F([Cdc2-Cyclin~{p1}]>0.15), 150). First values found :

parameter(k3,10). parameter(k4,120).





biocham: search_parameters([k3,k4],[(0,200),(0,200)],20,

period(Cdc2-Cyclin~{p1},35), 150).

First values found : parameter(k3,10). parameter(k4,180).





Parameter Optimization Evolutionary Algorithm

- Objective function: continuous satisfaction degree of LTL(R) formulae
- Covariance Matrix Adaptive Evolution Strategy (CMA-ES) [Hansen 01]











Generation 6





Landscape for Oscillation Amplitude Constraint

$\phi^*:$ F([A]>x \wedge F([A]<y)); amplitude x-y \geq 0.3



Bifurcation diagram

LTL satisfaction diagram

MPCE 2018



Unexpected Behavior of NIH3T3 Fibroblasts: Acceleration of the Clock at high FBS !

Time series data in individual mice fibroblasts [Feillet Delaunay 2012] Fluorescent markers of the cell cycle and the circadian clock (RevErb α) Medium with various concentrations of serum (FBS)

- FBS modulates the cell cycle frequency
- No observed time gating for mitosis
- But observed acceleration of the circadian clock in fastly dividing cells ! and not in confluent cells (24h)
 FBS 10% → Cell cycle 22h → Circadian clock 22h, phase 7h FBS 15% → Cell cycle 19h → Circadian clock 18h, phase 7h
 Statistical model phase locking [Feillet et al Delaunay Rand PNAS 2014]







Reverse Effect Cell Cycle \rightarrow Clock



Mechanistic model for this reverse effect ?

Hypothesis 1: Uniform inhibition of gene transcription during mitosis

- Entrainment in period
- No parameter values for correct entrainment in phase

Hypothesis 2: Selective regulation of clock genes during mitosis

- Entrainment in period and phase fitted to experimental data
- Prediction of reverb up-regulation during mitosis (or Bmal1 down)

[Traynard, Feillet, Soliman, Delaunay, F., Biosystems 2016]

Relogio-Herzel Model of the Circadian Clock (2011)

- 20 species, 71 parameters
- 60 parameters fitted to liver cell data
 - amplitude, period and phase data
- Per, Cry, Reverb, Ror, Bmal genes

Relógio, A., Westermark, P. O., Wallach, T., Schellenberg, K., Kramer, A., & Herzel, H. (2011). Tuning the mammalian circadian clock: robust synergy of two loops. PLoS Computational Biology.







Hypothesis 1: Uniform Inhibition of Transcription during Mitosis [Kang et al. 2008]



- Correct acceleration of both the cell cycle and the circadian clock
- But impossible to fit experimental phase shift between cell division time and RevErb peak
 - Experimental phase: 7h
 - Model phase: 18h



MPF
 RevErb::nucl
 Bmal-Clock::nucl
 Cry-Per::nucl



Hypothesis 2: Selective Regulation of Clock Genes during Mitosis



- Correct fit to period and phase experimental data (playing with only coupling strength regulation parameters)
- Two sets of parameter values fit the data:

	Parameters	First set	Second set
either down-regulation of Bmal1 or up-regulation of RevErb α during mitosis	Synthesis coefficient for Per	0.66	2.40
	Synthesis coefficient for Cry	2.30	0.67
	Synthesis coefficient for $RevErb-\alpha$	1.04	1.92
	Synthesis coefficient for Ror	2.1	1.51
	Synthesis coefficient for <i>Bmal1</i>	0	0.78
	Duration	2.97h	$2.81\mathrm{h}$

Hypothesis 2: Predictions



Prediction: different behaviors for a slow cell cycle (5% FBS)

3.5 3 2.5 timeinhib 2 1.5 Stronger 1 control of 0.5 the clock by the divisions 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22 **MPCE 2018** Faster cell cycle kdie

Score for the property: The cell cycle and the circadian clock have the same period

1

0.8

0.6

0.4

0.2

Complex Behaviors with High Variability observed after Treatment by Dexamethasone

• Dexamethasone synchronize cellular clocks, but complex dynamics observed

Medium	Clock period	Division period	Mean delay
FBS 10%	24.2 h \pm 0.5 h	$20.1~\mathrm{h}\pm0.94~\mathrm{h}$	10.7 h
FBS 20%	21.25 h ± 0.36 h	19.5 h ± 0.42 h	8.3 h
	$29~\mathrm{h}{\pm}1.05~\mathrm{h}$	16.05 h±0.48 h	6h/12h/22h



Interpreted as 5:4 and 1:1 locking modes for 10% FBS and 3:2 and 1:1 for 15%

[C. Feillet et al. Phase locking and multiple oscillating attractors for the coupled mammalian clock and cell cycle., PNAS 2014]

In our model, Dex pulse modeled by induction of high level of Per.
 Clock perturbation varies according to the time T of the pulse
 Stabilization of the clock may occur after 70h beyond observed data...

peak-peak distance in [18.8, 22.7] with T=162h [20.9, 21.7] with T=170h



Wrap-up

- Bi-directional coupling of cell cycle and circadian clock through
 - 1. Regulation of cell cycle genes (Wee1, p21, Myc) by clock genes (Bmal1,Per,Rev)
 - 2. Regulation of clock genes by cell cycle (up regulation of Rev-Erb α during mitosis)
- Modeling of molecular interaction networks by formal reaction systems
 - Differential (ODE), stochastic (CTMC), Petri Net (PN), Boolean semantics
 - Before simulation, static analysis and model-checking methods
- Modeling of temporal behaviors
 - quantitative temporal logic language FO-LTL
 - Sensitivity&robustness measures, parameter search w.r.t. FO-LTL specification
- BIOCHAM free software for modeling
 - reaction and influence systems (SBML compatible)
 - analysis and synthesis
- Same parameter search methods can be used for model-based control
 - Drug exposure optimization
 - External control of cell processes
 - Cell reprogramming

