



HAL
open science

Machinerie cellulaire et programmation biochimique: vers une informatique de la cellule

François Fages

► **To cite this version:**

François Fages. Machinerie cellulaire et programmation biochimique: vers une informatique de la cellule. Forum des lauréats en informatique et mathématiques appliquées 2014, 2014, Collège de France, Paris, France. hal-01103343

HAL Id: hal-01103343

<https://inria.hal.science/hal-01103343>

Submitted on 14 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Machinerie cellulaire et programmation biochimique: vers une informatique de la cellule



François Fages
Inria Paris-Rocquencourt
<http://lifeware.inria.fr/>



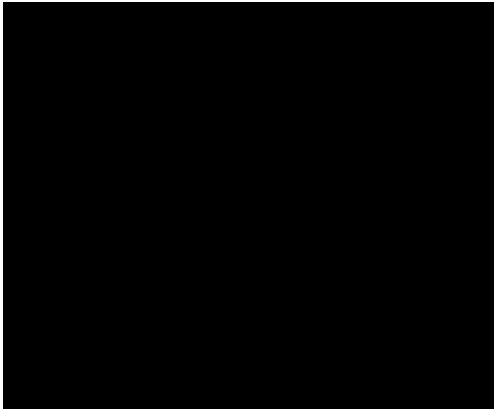
Objectif: **maîtriser la complexité des processus cellulaires** par

- Concepts de théorie de la programmation
- Méthodes de vérification de circuits et de programmes
- Algorithmes de recherche et d'optimisation

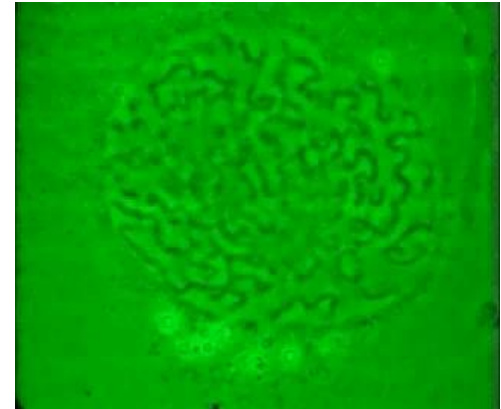
Logiciel de modélisation BIOCHAM v3.6 (machine abstraite biochimique)

Contexte: Biologie des Systèmes

“Gain system-level understanding of multi-scale biological processes in terms of their elementary interactions at the molecular level.” [Kitano 1999]



vidéos de divisions cellulaires et mitoses
[Lodish et al. 03]



Suite du projet de séquençage du génome humain (2000)

- Bases de données post-génomiques : interactions [protéines-protéines](#), [ARNs](#), ...
- Création d'[entrepôts de modèles](#) de processus cellulaires : [BioModels.net](#)
- [SBML](#) format d'échange de modèles
- [Logiciels de modélisation](#) (Cell designer, Cytoscape, Copasi, [BIOCHAM](#), [BIOPLM](#),...)
- Simulation d'une cellule entière *mycoplasma genitalium* [Karr Covert et al 12]

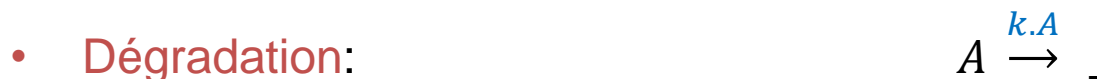
Programmer avec des réactions biochimiques ?

- Liaison, complexation: $A + B \rightarrow C$
 $cdk1 + cycB \rightarrow cdk1cycB$
- Déliaison, décomplexation: $A \rightarrow B + C$
- Modification, phosphorylation, transport: $A \rightarrow B$
 $cdk1cycB \rightarrow cdk1cycBp$
- Synthèse ARN, protéines: $A \rightarrow A + B$
 $E2Fb \rightarrow E2Fb + RNAcycB$
- Dégradation: $A \rightarrow _$

Programmer avec des réactions biochimiques ?



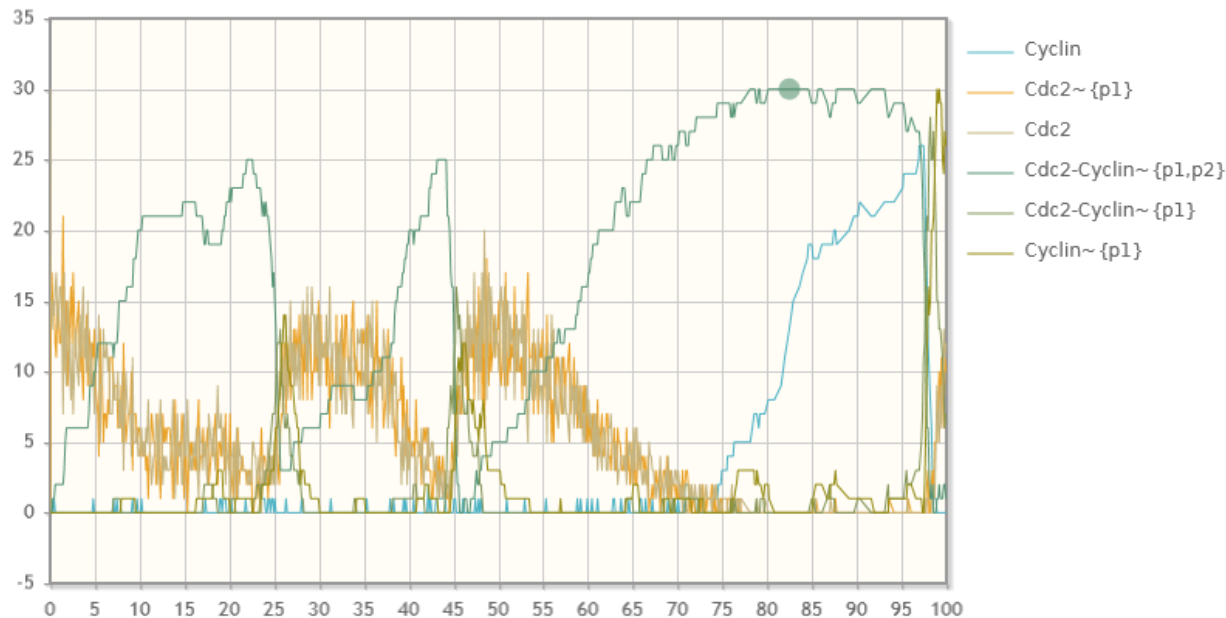
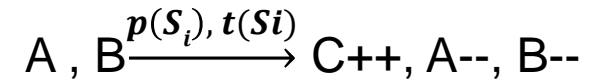
Le temps compte



Sémantiques des réactions ? $A+B \xrightarrow{f(A,B)} C$

Sémantique stochastique: nombres de molécules, probabilité et temps de transition

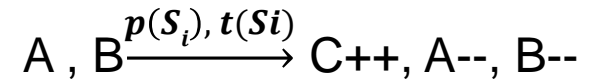
Chaîne de Markov à temps continu (CTMC)



Sémantiques des réactions ? $A+B \xrightarrow{f(A,B)} C$

Sémantique stochastique: nombres de molécules, probabilité et temps de transition

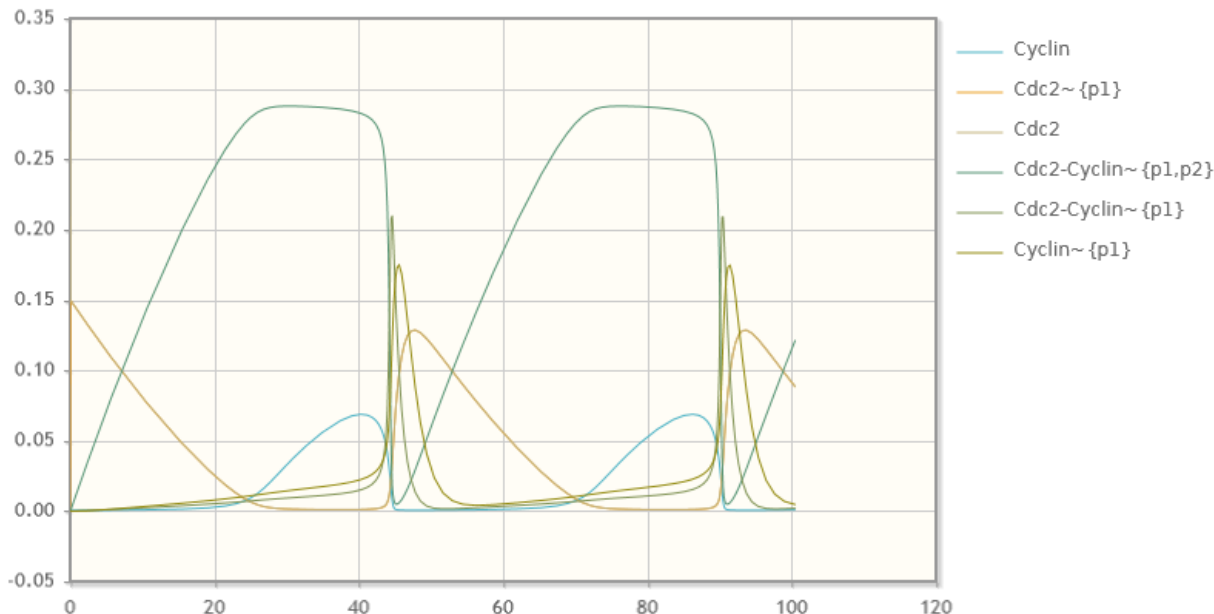
Chaîne de Markov à temps continu (CTMC)



Sémantique continue: concentrations, évolution continue

Equation différentielle ordinaire (EDO)

$$\frac{dA_i}{dt} = \sum_{r=1}^n f_r \times \delta_r(A_i)$$



Sémantiques des réactions ? $A+B \xrightarrow{f(A,B)} C$

Sémantique stochastique: nombres de molécules, probabilité et temps de transition

Chaîne de Markov à temps continu (CTMC) $A, B \xrightarrow{p(S_i), t(S_i)} C_{++}, A_{--}, B_{--}$

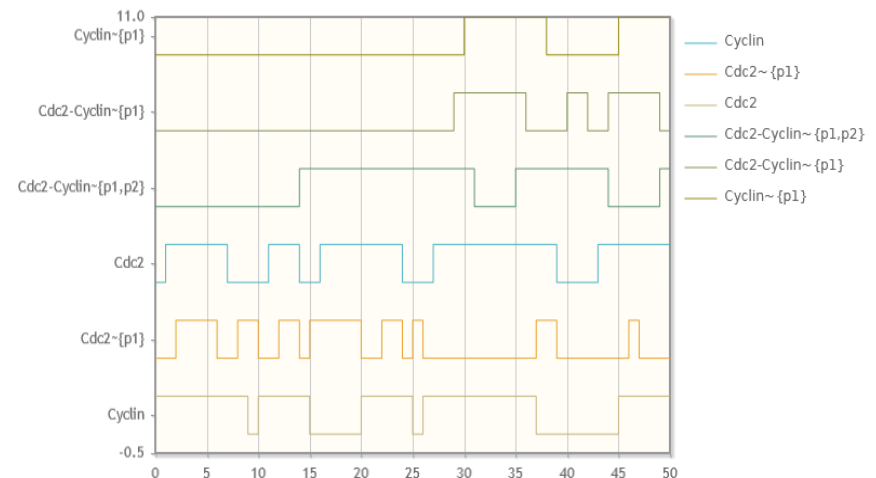
Sémantique continue: concentrations, évolution continue

Equation différentielle ordinaire (EDO) $\frac{dA_i}{dt} = \sum_{r=1}^n f_r \times \delta_r(A_i)$

Sémantique réseau de Petri: nombres de molécules $A, B \rightarrow C_{++}, A_{--}, B_{--}$

Réécriture de multiensembles

CHAM [Berry Boudol 90] [Banatre Le Metayer 86]



Sémantiques des réactions ? $A+B \xrightarrow{f(A,B)} C$

Sémantique stochastique: nombres de molécules, probabilité et temps de transition

Chaîne de Markov à temps continu (CTMC) $A, B \xrightarrow{p(S_i), t(S_i)} C_{++}, A_{--}, B_{--}$

Sémantique continue: concentrations, évolution continue

Equation différentielle ordinaire (EDO) $\frac{dA_i}{dt} = \sum_{r=1}^n f_r \times \delta_r(A_i)$

Sémantique réseau de Petri: nombres de molécules $A, B \rightarrow C_{++}, A_{--}, B_{--}$

Réécriture de multiensembles

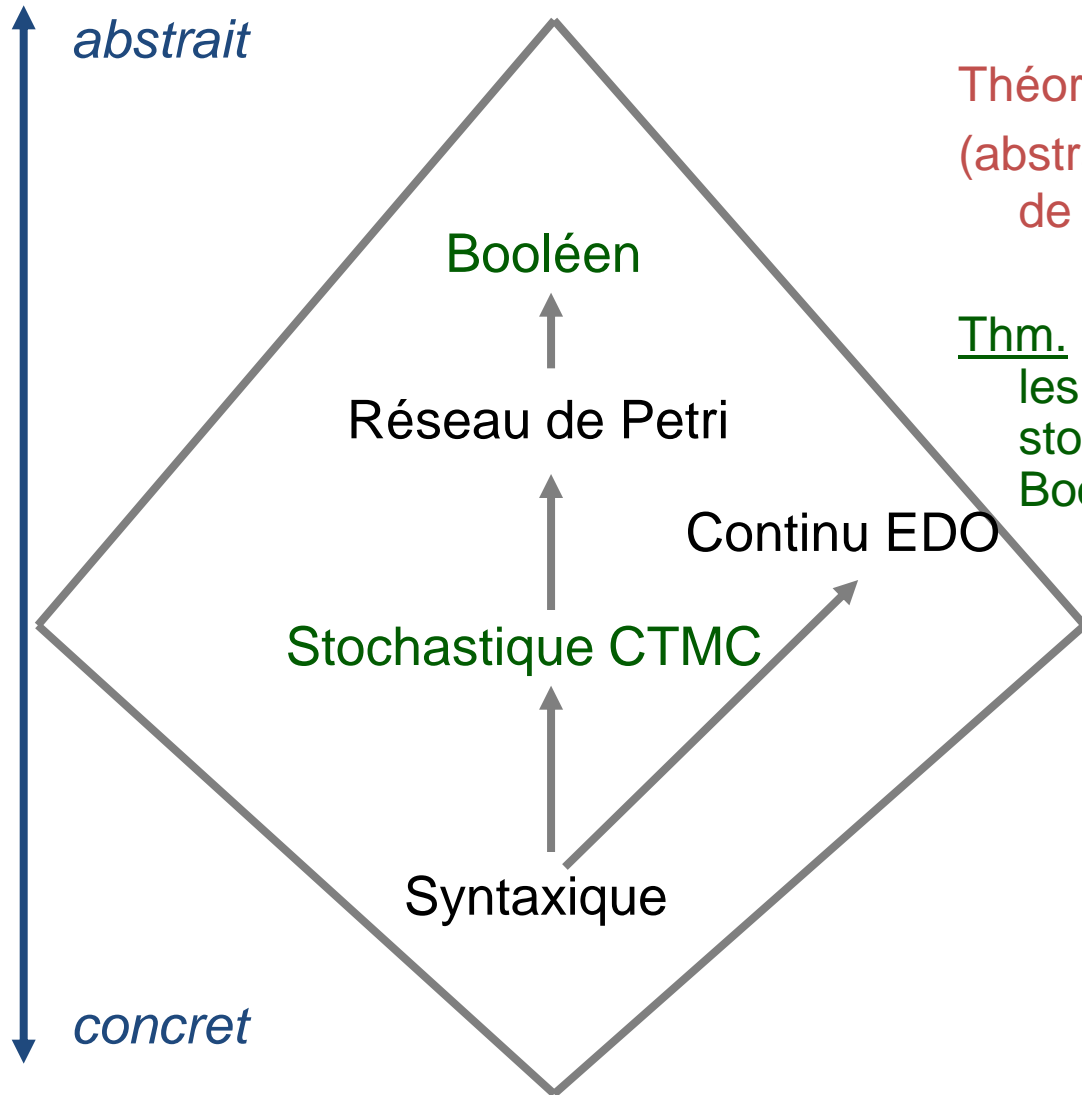
CHAM [Berry Boudol 90] [Banatre Le Metayer 86]

Sémantique booléenne: présence-absence

Système de transition asynchrone

$A \wedge B \rightarrow C \wedge A/\neg A \wedge B/\neg B$

Hiérarchie de sémantiques



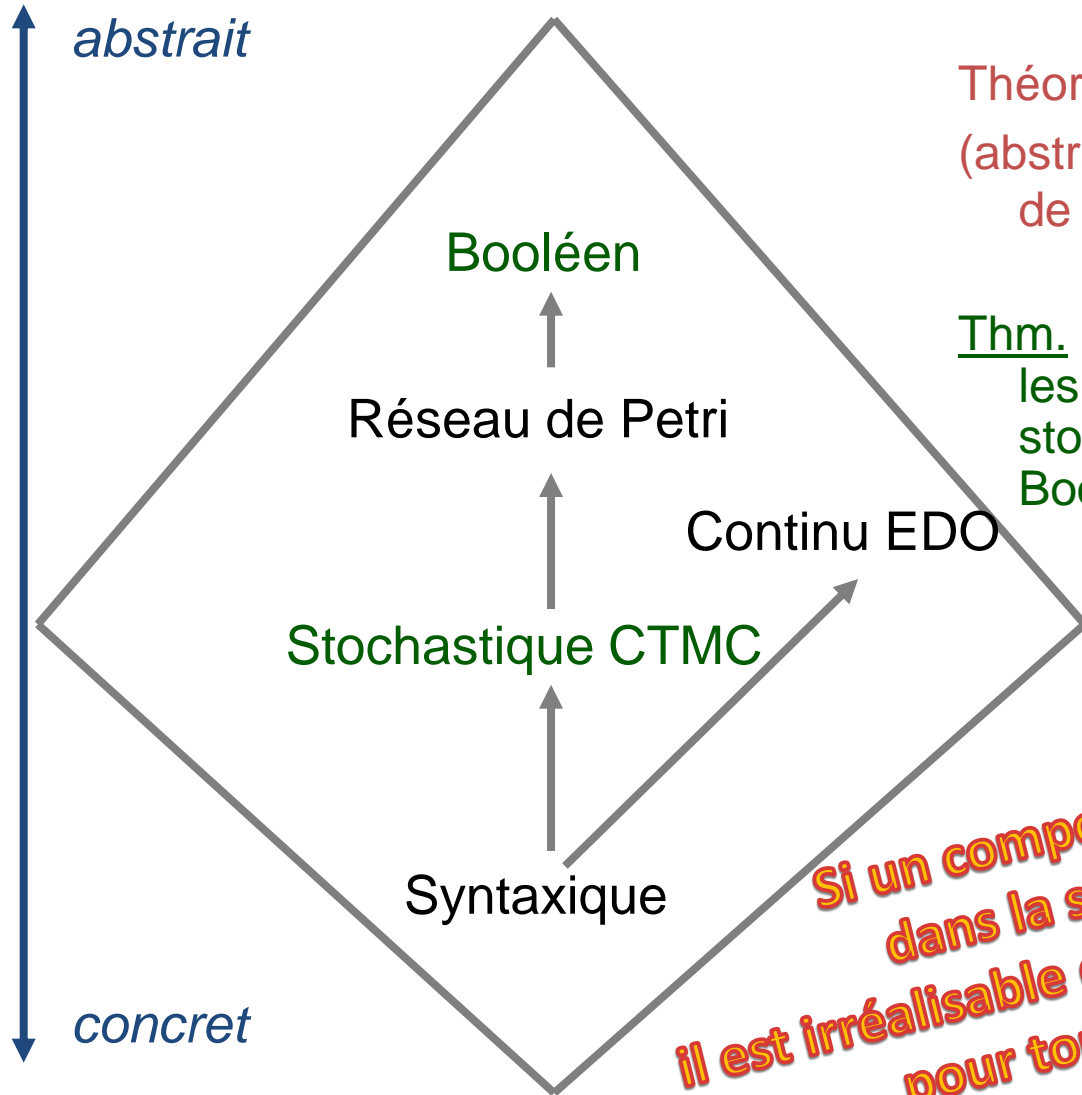
Théorie de l'interprétation abstraite
(abstractions comme correspondances
de Galois entre treillis d'information)

[Cousot Cousot POPL'77]

Thm. Correspondances de Galois entre
les sémantiques de traces
stochastiques, réseaux de Petri et
Booléenne des systèmes de réactions

[FF Soliman CMSB'06, TCS'08]

Hiérarchie de sémantiques



Théorie de l'interprétation abstraite
(abstractions comme correspondances
de Galois entre treillis d'information)

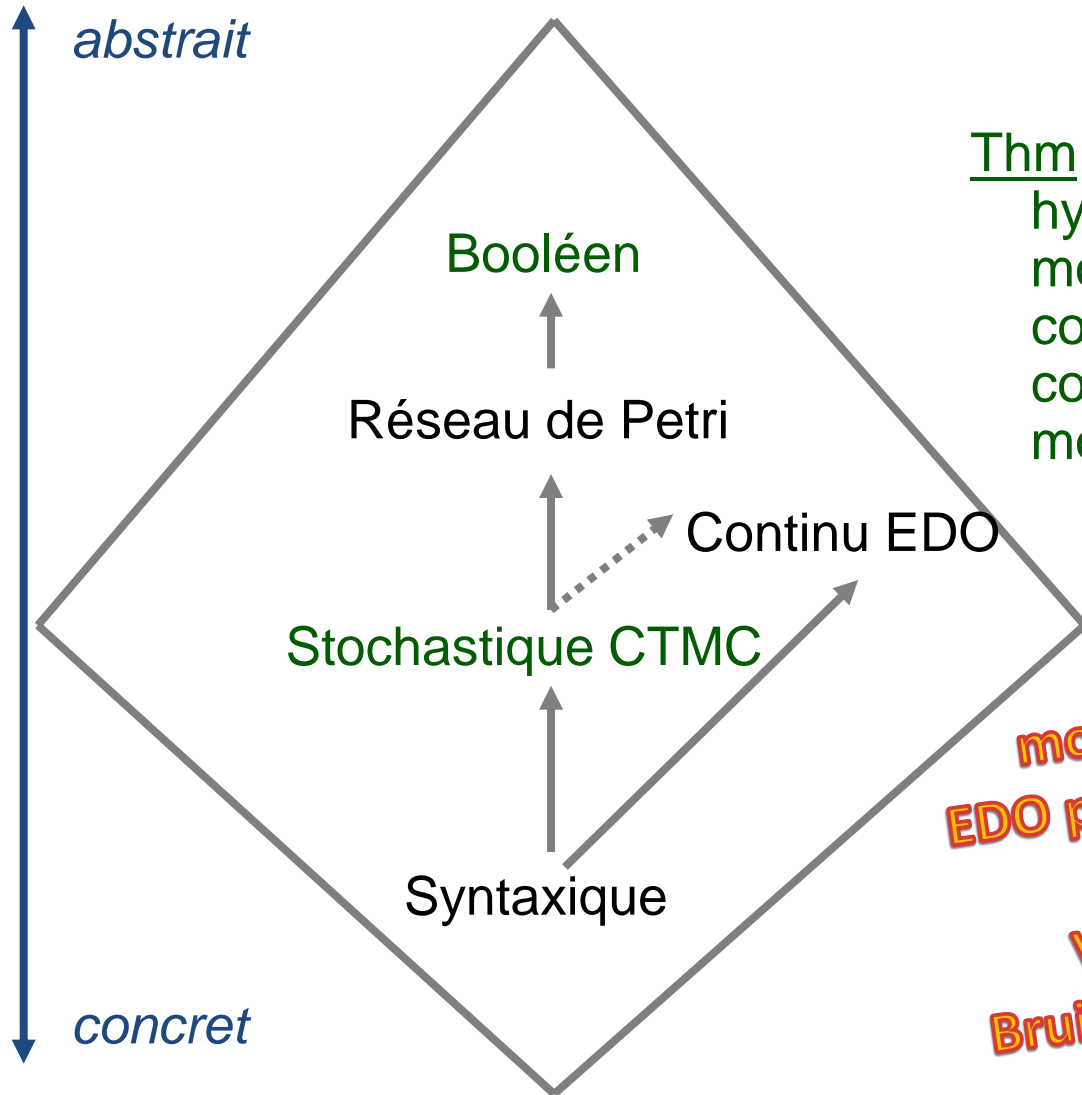
[Cousot Cousot POPL'77]

Thm. Correspondances de Galois entre
les sémantiques de traces
stochastiques, réseaux de Petri et
Booléenne des systèmes de réactions

[FF Soliman CMSB'06, TCS'08]

**Si un comportement est impossible
dans la sémantique booléenne
il est irréalisable dans la sémantique stochastique
pour toutes vitesses de réaction**

Hiérarchie de sémantiques

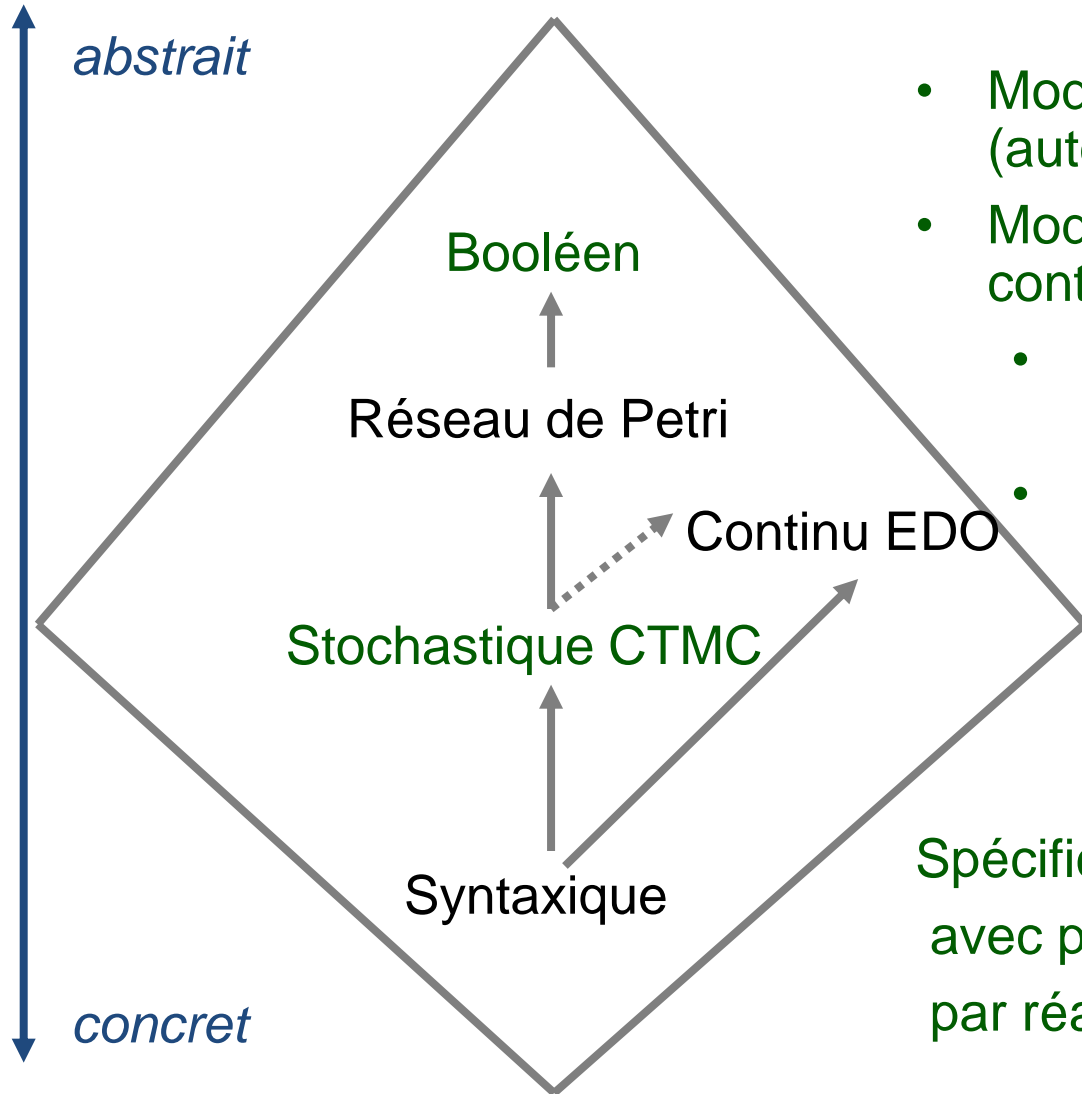


Thm. [Gillespie 71] Sous des hypothèses de grand nombre de molécules, la sémantique continue (EDO) approxime le comportement stochastique moyen (CTMC)

Sujet chaud:
moments d'ordre supérieur
EDO pour moyenne et variance,...

Variabilités entre cellules
Bruits intrinsèque et extrinsèque

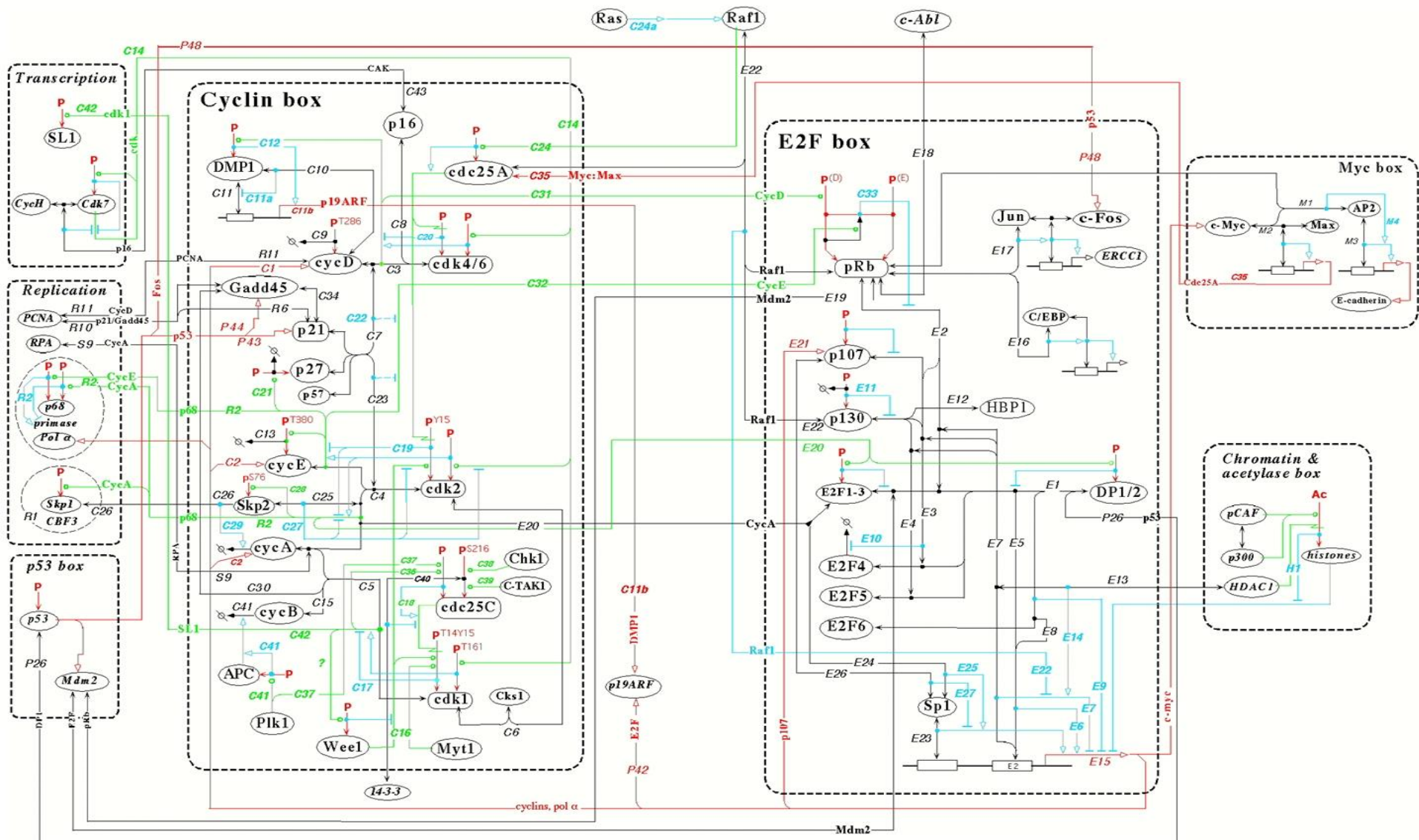
Modèles hybrides



- Modèles hybrides booléen-continu (automates hybrides)
- Modèles hybrides stochastiques-continus (CTMC+ODE)
 - Expression stochastique des gènes
 - activation continue des protéines

Spécification de simulateurs hybrides avec partitionnement dynamique par réactions et événements SBML

Carte de commande du cycle cellulaire [Kohn 99]



Model-checking symbolique de la carte de Kohn

800 réactions, 165 protéines and gènes, 500 variables, 2^{500} states.

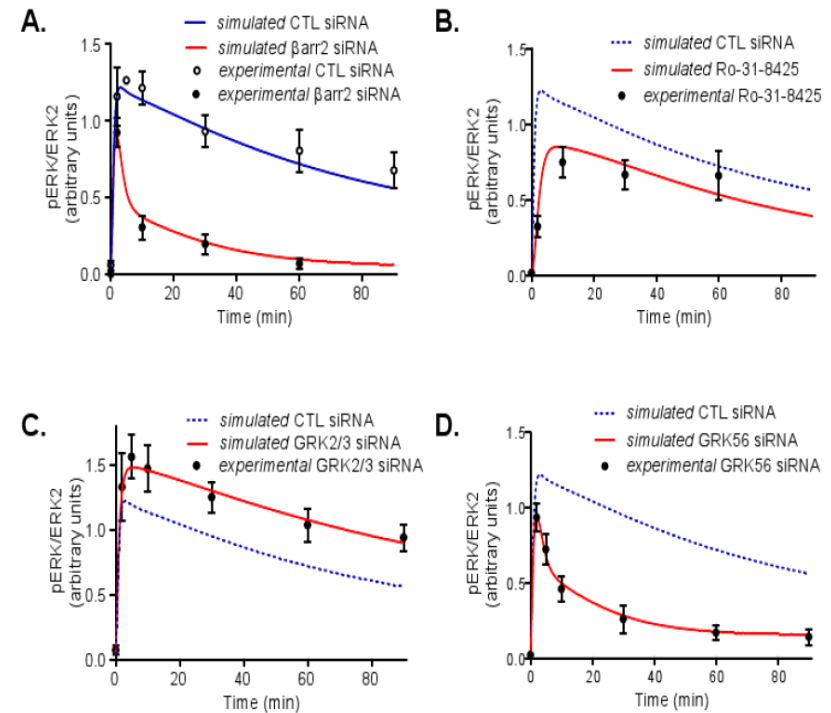
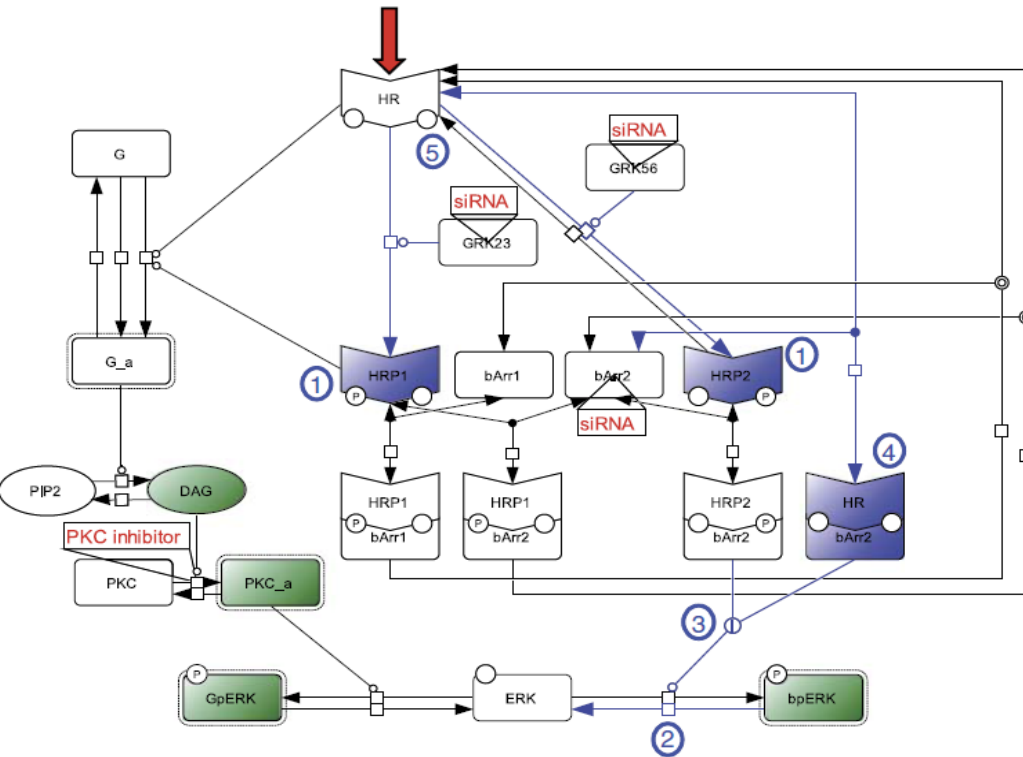
Temps en secondes du vérificateur NuSMV en BIOCHAM [Chabrier et al. TCS 2004]

| Initial state G2 | Query: | Time in sec. |
|-----------------------------------|--|--------------|
| | compiling | 29 |
| Reachability G1 | EF CycE | 2 |
| Reachability G1 | EF CycD | 1.9 |
| Reachability G1 | EF PCNA-CycD | 1.7 |
| Checkpoint for Mitosis complex | \neg EF (\neg Cdc25~{Nterm} U Cdk1~{Thr161}-CycB) | 2.2 |
| Oscillations CycA | EG ((EF \neg CycA) & (EF CycA)) | 31.8 |
| Oscillations CycB | EG ((EF \neg CycB) & (EF CycB)) | 6 |

Formalisation des comportements quantitatifs en logique temporelle

- **Formalisation des comportements dynamiques observés expérimentalement** :
 - Propriétés qualitatives: logique temporelle booléenne CTL
 - Propriétés quantitatives: logique avec contraintes numériques FO-LTL(R_{lin})
- **Spécifications formelles** pour la **validation de modèles**
 - Vérification symbolique booléenne [Chabrier Chiaverini Danos FF Schachter 04 TCS]
 - Vérification simulations numériques, robustesse [Rizk Batt FF Soliman 09 Bioinformatics]
- **Contraintes** pour la **synthèse de modèles** (inférence des paramètres cinétiques)
 - Algorithmes de recherche évolutionnaires, CMA-ES [Hansen 01, Rizk Batt FF Soliman 11 TCS]
 - Satisfaction → modèle prédictif quantitatif
 - Echec → réexamen de la structure du modèle → **contributions biologiques**

Succès en modélisation de la signalisation GPCR



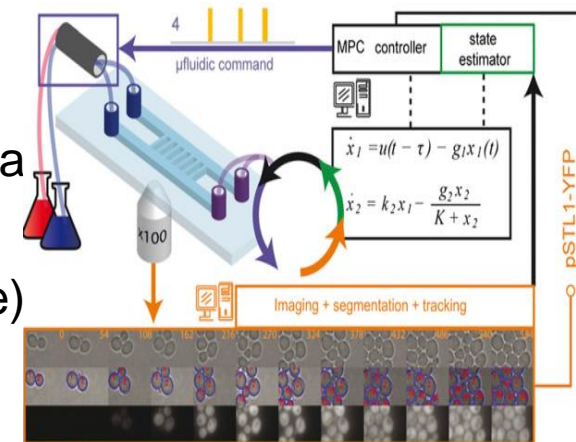
- Modèle réduit comprenant **4 observables**, **4 mutations**, interactions connues
- Echec des recherches de paramètres en **logique temporelle** en BIOCHAM
- **Révision de la structure** des interactions admises, vérifiée a posteriori
[Heitzler, ..., FF, Lefkowitz, Reiter 2012 *Molecular Systems Biology* 8(590)]

Commande de l'expression des gènes dans la levure

Boucle de **perception - apprentissage – action** sur 3 jours

1. Analyse d'image microscope (pistage de cellule ou popula
2. Inférence paramètres cinétiques du modèle (calibration)
3. Inférence paramètres de pression osmotique (commande)

[Uhlendorf, ..., Batt, Hersen *PNAS* 2012]



SM 5 : Real time yeast single cell control to a constant target profile

Long-term model predictive control of gene expression at the population and single-cell levels

J. Uhlendorf, A. Miermont, T. Delaveau, G. Charvin, F. Fages, S. Bottani, G. Batt* & P. Hersen*

✉ pascal.hersen@univ-paris-diderot.fr 🔗 <http://inryurl.com/TeamHersen>
 📧 gregory.batt@inria.fr 🔗 <http://www-roq.inria.fr/~batt/>



SM 7 : Real time yeast single cell control to a trapeze target profile

Long-term model predictive control of gene expression at the population and single-cell levels

J. Uhlendorf, A. Miermont, T. Delaveau, G. Charvin, F. Fages, S. Bottani, G. Batt* & P. Hersen*

✉ pascal.hersen@univ-paris-diderot.fr 🔗 <http://inryurl.com/TeamHersen>
 📧 gregory.batt@inria.fr 🔗 <http://www-roq.inria.fr/~batt/>



SM 6 : Real time yeast single cell control to a sine wave target profile

Long-term model predictive control of gene expression at the population and single-cell levels

J. Uhlendorf, A. Miermont, T. Delaveau, G. Charvin, F. Fages, S. Bottani, G. Batt* & P. Hersen*

✉ pascal.hersen@univ-paris-diderot.fr 🔗 <http://inryurl.com/TeamHersen>
 📧 gregory.batt@inria.fr 🔗 <http://www-roq.inria.fr/~batt/>



Résumé

- Théorie des réseaux de réactions biochimiques

- Hiérarchie de sémantiques
- Réductions de modèles
- Hypergraphe de réaction et graphe d'influence

*La simplicité est la sophistication suprême.
L. de Vinci*

- Formalisation des comportements en logique temporelle

- Vérification de modèles
- Inférence des valeurs des paramètres

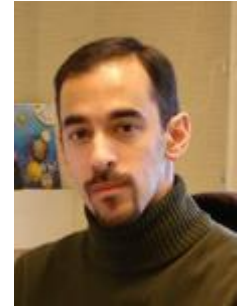
*Tous les modèles sont faux
mais certains sont utiles.
G.E.P. Box*

- Programmation biochimique

- Biologie synthétique: nouvelles fonctions dans des cellules ou vésicules non vivants
- Biologie des systèmes: élucidation des programmes naturels acquis par l'évolution

*What I cannot create I do not understand
R. Feynman*

Lifeware team at Inria



Research Scientists

Grégory Batt (Inria, Researcher, Synthetic Biology Principal Investigator)
Sylvain Soliman (Inria, Researcher)

Associates

Pascal Hersen (CNRS MSC lab, Researcher)
Denis Thieffry (ENS Paris, Professor)

PhD students

François Bertaux (AMX, Ecole Polytechnique, with EPI Bang)
Katherine Chiang (National Taiwan University)
Xavier Duportet (Inria, CORDI-S, MIT Weiss lab)
David Fournier (Cifre General Electric Transportation)
Steven Gay (Inria, Post doc Univ. Louvain la Neuve, Belgium)
Jean-Baptiste Lugagne (Inria, CORDI-C, CNRS MSC lab)
Artemis Llamosi (CNRS MSC lab, Inria)
Pauline Traynard (AMX, Ecole Polytechnique, with ENS)

Engineers

François-Marie Floch (Inria ADT, Engineer)
Thierry Martinez (Research Engineer, SED, Inria Rocquencourt)
Philippe Morignot (Inria, Engineer)

stagiaires
et
nouveaux
doctorants
bienvenus !

