Activity Report 2017

Project-Team LIFEWARE

Computational systems biology and optimization
# Table of contents

1. Personnel .................................................................................................................. 2
2. Overall Objectives ..................................................................................................... 2
3. Research Program ...................................................................................................... 3
   3.1. Computational Systems Biology 3
   3.2. Chemical Reaction Network (CRN) Theory 4
   3.3. Logical Paradigm for Systems Biology 4
   3.4. Modeling of Phenotypic Heterogeneity in Cellular Processes 4
   3.5. External Control of Cell Processes 5
   3.6. Synthesis of Continuous CRNs 5
   3.7. Constraint Solving and Optimization 6
4. Application Domains .................................................................................................. 6
   4.1. Preamble 6
   4.2. Modeling software for systems biology 6
   4.3. Couplings between the cell cycle and the circadian clock 7
   4.4. Biosensor design and implementation in non-living protocells 7
5. Highlights of the Year ............................................................................................... 7
6. New Software and Platforms .................................................................................... 8
7. New Results ................................................................................................................ 8
   7.1. Strong Turing completeness of continuous CRNs 8
   7.2. Influence networks compared with reaction networks 8
   7.3. Machine learning influence networks from data 9
   7.4. Shaping bacterial population behavior through computer-interfaced control of individual cells 9
   7.5. Balancing a genetic toggle switch by real-time feedback control and periodic forcing 9
   7.6. Abstracting the dynamics of biological pathways using information theory: a case study of apoptosis pathway 10
   7.7. Long-term tracking of budding yeast cells in brightfield microscopy: CellStar and the Evaluation Platform 10
   7.8. Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability 11
   7.9. Recombinase-based genetic circuit optimization by integer linear programming 11
   7.10. Coupled models of the cell cycle and circadian clock 11
8. Partnerships and Cooperations ................................................................................... 11
   8.1. National Initiatives 11
   8.1.1. ANR Projects 11
   8.1.2. Inria Project Lab 12
   8.2. European Initiatives 12
   8.3. International Initiatives 12
9. Dissemination ............................................................................................................ 12
   9.1. Promoting Scientific Activities 12
   9.1.1. Scientific Events Organisation 12
   9.1.2. Scientific Events Selection 13
   9.1.2.1. Member of the Conference Program Committees 13
   9.1.2.2. Reviewer 13
   9.1.3. Journal 13
   9.1.3.1. Member of the Editorial Boards 13
   9.1.3.2. Reviewer - Reviewing Activities 13
   9.1.4. Invited Talks 13
   9.1.5. Leadership within the Scientific Community 14
   9.1.6. Scientific Expertise 15
9.1.7. Research Administration 15
9.2. Teaching - Supervision - Juries 15
  9.2.1. Teaching 15
  9.2.2. Supervision 16
  9.2.3. Juries 16
9.3. Popularization 16
10. Bibliography ................................................................. 16
Project-Team LIFEWARE

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 April 01

Keywords:

**Computer Science and Digital Science:**
- A2.1.1. - Semantics of programming languages
- A2.1.5. - Constraint programming
- A2.1.10. - Domain-specific languages
- A2.2.1. - Static analysis
- A2.3.2. - Cyber-physical systems
- A2.4. - Verification, reliability, certification
- A2.4.1. - Analysis
- A2.4.2. - Model-checking
- A2.4.3. - Proofs
- A6.1.1. - Continuous Modeling (PDE, ODE)
- A6.1.2. - Stochastic Modeling (SPDE, SDE)
- A6.1.3. - Discrete Modeling (multi-agent, people centered)
- A6.1.4. - Multiscale modeling
- A6.2.4. - Statistical methods
- A6.2.6. - Optimization
- A6.3.1. - Inverse problems
- A6.3.4. - Model reduction
- A7.2. - Logic in Computer Science
- A8.1. - Discrete mathematics, combinatorics
- A8.2. - Optimization
- A8.7. - Graph theory
- A9.7. - AI algorithmics

**Other Research Topics and Application Domains:**
- B1. - Life sciences
- B1.1.2. - Molecular biology
- B1.1.3. - Cellular biology
- B1.1.9. - Bioinformatics
- B1.1.10. - Mathematical biology
- B1.1.11. - Systems biology
- B1.1.12. - Synthetic biology
- B1.1.14. - Microbiology
- B2. - Health
- B2.2.3. - Cancer
- B2.4.1. - Pharmaco kinetics and dynamics
- B2.4.2. - Drug resistance
- B9. - Society and Knowledge
1. Personnel

Research Scientists
- François Fages [Team leader, Inria, Senior Researcher, HDR]
- Grégory Batt [Inria, Senior Researcher, HDR]
- Jakob Ruess [Inria, Researcher]
- Sylvain Soliman [Inria, Researcher, HDR]

Faculty Member
- Philippe Dague [Univ Paris-Sud, Professor, from Feb 2017 until Aug 2017]

Post-Doctoral Fellows
- Jean-Baptiste Lugagne [Inria, until Jan 2017, then MSC lab (CNRS/Paris7) until Aug 2017]
- Chiara Fracassi [Inria, until Aug 2017]

PhD Students
- Chetan Aditya [Inria, from Nov 2017]
- Virgile Andreani [Ecole Polytechnique]
- Jean Baptiste Caron [Inria, until Feb 2017]
- Sebastian Ramon Sosa Carrillo [Inria, from Aug 2017]

Technical staff
- David Coudrin [Inria, from Nov 2017]
- Steven Fletcher [Inria, from Apr 2017]

Interns
- Sebastian Ramon Sosa Carrillo [Inria, AIV master, until Apr 2017]
- Arthur Carcano [Inria, master MPRI, Ecole Normale Supérieure Paris, from Mar 2017 to Jul 2017]
- Elea Greugny [Pasteur, 4th year student, INSA Lyon, from Apr 2017 to Aug 2017]
- Jeremy Grignard [Inria, from Apr 2017 to Aug 2017]

Administrative Assistant
- Corinne Petitot [Inria]

External Collaborator
- Denis Thieffry [Ecole Normale Supérieure Paris, HDR]

2. Overall Objectives

2.1. Overall Objectives

This project aims at developing formal methods and experimental settings for understanding the cell machinery and establishing formal paradigms in cell biology. It is based on the vision of cells as machines, biochemical reaction systems as programs, and on the use of concepts and tools from computer science to master the complexity of cell processes. While for the biologist, as well as for the mathematician, the size of the biological networks and the number of elementary interactions constitute a complexity barrier, for the computer scientist the difficulty is not that much in the size of the networks than in the unconventional nature of biochemical computation. Unlike most programs, biochemical reaction systems involve transitions that are stochastic rather than deterministic, continuous-time rather than discrete-time, poorly localized in compartments instead of well-structured in modules, and created by evolution instead of by rational design. It is our belief however that some form of modularity is required by an evolutionary system to survive, and that the elucidation of these modules in biochemical computation is now a key to apply engineering methods in cell biology on a large scale.
Concretely, we keep developing a theory of biochemical computation with a prototype implementation in the Biochemical Abstract Machine BIOCHAM, a modeling and analysis platform for Systems Biology. The reaction rule-based language used in this system allows us to reason about biochemical reaction networks at different levels of abstraction, in either the stochastic, differential, discrete, logical or hybrid semantics of the reactions. This makes it possible to apply a variety of static analysis methods, before going to simulations and dynamical analyses, for which we use temporal logics as a specification language to formalize biological behaviours with imprecise data, validate models w.r.t. observations, constrain model building, and calibrate models in high dimension by optimization methods.

A tight integration between dry lab and wet lab efforts is also essential for the success of the project. In collaboration with biologists, we investigate concrete biological questions and develop computational models fitted to quantitative data which allow us to make quantitative predictions. In collaboration with Pascal Hersen, MSC lab, we contribute to the development of an experimental platform for the closed-loop control of intracellular processes. This platform combines hardware (microfluidic device and microscope), software (cell tracking and model-based predictive control algorithms) and genetically modified living cells. It is used to investigate the possibilities to externalize the control of intracellular processes for systems and synthetic biology applications, and perform accurate observations, modifications and real-time control at both single cell and cell population levels. We are affiliated with the Doctorate Schools “Frontières du vivant (FdV)” of University Sorbonne Paris Cité and “Sciences et technologies de l’information et de la communication (STIC)” of University Paris-Saclay.

This project addresses fundamental research issues in computer science on the interplay between structure and dynamics in large interaction networks, and on the mixing of continuous and discrete computation. Many static analysis problems of biological networks are NP-hard. The recourse to constraint logic programming (CLP) to model and solve them, is our secret weapon, which probably explains our capability to experiment ideas in computational systems biology in very short time, by implementing them in CLP, integrating them as new components in our modeling platform BIOCHAM, and evaluating them directly on a large scale in systems biology model repositories such as BIOMODELS.NET.

3. Research Program

3.1. Computational Systems Biology

Bridging the gap between the complexity of biological systems and our capacity to model and quantitatively predict system behaviors is a central challenge in systems biology. We believe that a deeper understanding of the concept and theory of biochemical computation is necessary to tackle that challenge. Progress in the theory is necessary for scaling, and enabling the application of static analysis, module identification and decomposition, model reductions, parameter search, and model inference methods to large biochemical reaction systems. A measure of success on this route will be the production of better computational modeling tools for elucidating the complex dynamics of natural biological processes, designing synthetic biological circuits and biosensors, developing novel therapy strategies, and optimizing patient-tailored therapeutics.

Progress on the coupling of models to data is also necessary. Our approach based on quantitative temporal logics provides a powerful framework for formalizing experimental observations and using them as formal specification in model building. Key to success is a tight integration between in vivo and in silico work, and on the mixing of dry and wet experiments, enabled by novel biotechnologies. In particular, the use of microfluidic devices makes it possible to measure behaviors at both single-cell and cell population levels in vivo, provided innovative modeling, analysis and control methods are deployed in silico.

In synthetic biology, while the construction of simple intracellular circuits has shown feasible, the design of larger, multicellular systems is a major open issue. In engineered tissues for example, the behavior results from the subtle interplay between intracellular processes (signal transduction, gene expression) and intercellular processes (contact inhibition, gradient of diffusible molecule), and the question is how should cells be genetically modified such that the desired behavior robustly emerges from cell interactions.
3.2. Chemical Reaction Network (CRN) Theory

Feinberg’s chemical reaction network theory and Thomas’s influence network analyses provide sufficient and/or necessary structural conditions for the existence of multiple steady states and oscillations in regulatory networks. Those conditions can be verified by static analyzers without knowing kinetic parameter values nor making any simulation. In this domain, most of our work consists in analyzing the interplay between the structure (Petri net properties, influence graph, subgraph epimorphisms) and the dynamics (Boolean, CTMC, ODE, time scale separations) of biochemical reaction systems. In particular, our study of influence graphs of reaction systems, our generalization of Thomas’ conditions of multi-stationarity and Soulé’s proof to reaction systems \(^1\), the inference of reaction systems from ODEs \(^2\), the computation of structural invariants by constraint programming techniques, and the analysis of model reductions by subgraph epimorphisms now provide solid ground for developing static analyzers, using them on a large scale in systems biology, and elucidating modules.

3.3. Logical Paradigm for Systems Biology

Our group was among the first ones in 2002 to apply model-checking methods to systems biology in order to reason on large molecular interaction networks, such as Kohn’s map of the mammalian cell cycle (800 reactions over 500 molecules) \(^3\). The logical paradigm for systems biology that we have subsequently developed for quantitative models can be summarized by the following identifications:

- Biological model = transition system \(K\)
- Dynamical behavior specification = temporal logic formula \(\phi\)
- Model validation = model-checking \(K\), \(s \models \phi\)
- Model reduction = sub-model-checking, \(K' \subset K\) s.t. \(K'\models \phi\)
- Model prediction = formula enumeration, \(\phi\) s.t. \(K, s \models \phi\)
- Static experiment design = symbolic model-checking, state \(s\) s.t. \(K, s \models \phi\)
- Model synthesis = constraint solving \(K?\), \(s \models \phi\)
- Dynamic experiment design = constraint solving \(K?, s' \models \phi\)

In particular, the definition of a continuous satisfaction degree for first-order temporal logic formulae with constraints over the reals, was the key to generalize this approach to quantitative models, opening up the field of model-checking to model optimization \(^4\). This line of research continues with the development of temporal logic patterns with efficient constraint solvers and their generalization to handle stochastic effects.

3.4. Modeling of Phenotypic Heterogeneity in Cellular Processes

Since nearly two decades, a significant interest has grown for getting a quantitative understanding of the functioning of biological systems at the cellular level. Given their complexity, proposing a model accounting for the observed cell responses, or better, predicting novel behaviors, is now regarded as an essential step to validate a proposed mechanism in systems biology. Moreover, the constant improvement of stimulation and observation tools creates a strong push for the development of methods that provide predictions that are increasingly precise (single cell precision) and robust (complex stimulation profiles).

---


\(^4\) On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology A. Rizk, G. Batt, F. Fages, S. Soliman International Conference on Computational Methods in Systems Biology, 251-268
It is now fully apparent that cells do not respond identically to a same stimulation, even when they are all genetically-identical. This phenotypic heterogeneity plays a significant role in a number of problems ranging from cell resistance to anticancer drug treatments to stress adaptation and bet hedging.

Dedicated modeling frameworks, notably stochastic modeling frameworks, such as chemical master equations, and statistic modeling frameworks, such as ensemble models, are then needed to capture biological variability.

Appropriate mathematical and computational should then be employed for the analysis of these models and their calibration to experimental data. One can notably mention global optimization tools to search for appropriate parameters within large spaces, moment closure approaches to efficiently approximate stochastic models, and (stochastic approximations of) the expectation maximization algorithm for the identification of mixed-effects models.

3.5. External Control of Cell Processes

External control has been employed since many years to regulate culture growth and other physiological properties. Recently, taking inspiration from developments in synthetic biology, closed loop control has been applied to the regulation of intracellular processes. Such approaches offer unprecedented opportunities to investigate how a cell process dynamical information by maintaining it around specific operating points or driving it out of its standard operating conditions. They can also be used to complement and help the development of synthetic biology through the creation of hybrid systems resulting from the interconnection of in vivo and in silico computing devices.

In collaboration with Pascal Hersen (CNRS MSC lab), we developed a platform for gene expression control that enables to control protein concentrations in yeast cells. This platform integrates microfluidic devices enabling long-term observation and rapid change of the cells environment, microscopy for single cell measurements, and software for real-time signal quantification and model based control. We demonstrated in 2012 that this platform enables controlling the level of a fluorescent protein in cells with unprecedented accuracy and for many cell generations.

More recently, motivated by an analogy with a benchmark control problem, the stabilization of an inverted pendulum, we investigated the possibility to balance a genetic toggle switch in the vicinity of its unstable equilibrium configuration. We searched for solutions to balance an individual cell and even an entire population of heterogeneous cells, each harboring a toggle switch.

Independently, in collaboration with colleagues from IST Austria, we investigated the problem of controlling cells, one at a time, by constructing an integrated optogenetic platform. It enables experiments that bridge individual and population behaviors. We demonstrated: (i) population structuring by independent closed-loop control of gene expression in many individual cells, (ii) cell–cell variation control during antibiotic perturbation, (iii) hybrid bio-digital circuits in single cells, and freely specifiable digital communication between individual bacteria.

3.6. Synthesis of Continuous CRNs

The continuous nature of many protein interactions leads us to consider models of analog computation, and in particular, the recent results in the theory of analog computability and complexity obtained by Amaury

---


Pouly and Olivier Bournez, establish fundamental links with digital computation. In [10], we derive from these results a Turing completeness result for elementary reaction systems (without polymerization) under the differential semantics. The proof of this result shows how mathematical functions described by Ordinary Differential Equations, namely by Polynomial Initial Value Problems (PIVP), can be compiled into elementary biochemical reactions, furthermore with a notion of analog computation complexity defined as the length of the trajectory to reach a given precision on the result. This opens a whole research avenue to analyze natural circuits in Systems Biology, transform behavioural specifications into biochemical reactions for Synthetic Biology, and compare artificial circuits with natural circuits acquired through evolution, from the novel point of view of analog computation complexity.

3.7. Constraint Solving and Optimization

Constraint solving and optimization methods are important in our research. On the one hand, static analysis of biochemical reaction networks involves solving hard combinatorial optimization problems, for which constraint programming techniques have shown particularly successful, often beating dedicated algorithms and allowing to solve large instances from model repositories. On the other hand, parameter search and model calibration problems involve similarly solving hard continuous optimization problems, for which evolutionary algorithms such as the covariance matrix evolution strategy (CMA-ES) has shown to provide best results in our context, for up to 100 parameters, for building challenging quantitative models, gaining model-based insights, revisiting admitted assumptions, and contributing to biological knowledge.

4. Application Domains

4.1. Preamble

Our collaborative work on biological applications is expected to serve as a basis for groundbreaking advances in cell functioning understanding, cell monitoring and control, and novel therapy design and optimization. Our collaborations with biologists are focused on concrete biological questions, and on the building of predictive models of biological systems to answer them. However, one important application of our research is the development of a modeling software for computational systems biology.

4.2. Modeling software for systems biology

Since 2002, we develop an open-source software environment for modeling and analyzing biochemical reaction systems. This software, called the Biochemical Abstract Machine (BIOCHAM), is compatible with SBML for importing and exporting models from repositories such as BioModels. It can perform a variety of static analyses, specify behaviors in Boolean or quantitative temporal logics, search parameter values satisfying temporal constraints, and make various simulations. While the primary reason of this development effort is to be able to implement our ideas and experiment them quickly on a large scale, BIOCHAM is used by other groups either for building models, for comparing techniques, or for teaching (see statistics in software section). BIOCHAM-WEB is a web application which makes it possible to use BIOCHAM without any installation. We plan to continue developing BIOCHAM for these different purposes and improve the software quality.

4.3. Couplings between the cell cycle and the circadian clock

Recent advances in cancer chronotherapy techniques support the evidence that there exist important links between the cell cycle and the circadian clock genes. One purpose for modeling these links is to better understand how to efficiently target malignant cells depending on the phase of the day and patient characteristics. These questions are at the heart of our collaboration with Franck Delaunay (CNRS Nice) and Francis Lévi (Univ. Warwick, GB, formerly INSERM Hopital Paul Brousse, Villejuif) and of our participation in the ANR HYCLOCK project and in the submitted EU H2020 C2SyM proposal, following the former EU EraNet Sysbio C5SYS and FP6 TEMPO projects. In the past, we developed a coupled model of the Cell Cycle, Circadian Clock, DNA Repair System, Irinotecan Metabolism and Exposure Control under Temporal Logic Constraints. We now focus on the bidirectional coupling between the cell cycle and the circadian clock and expect to gain fundamental insights on this complex coupling from computational modeling and single-cell experiments.

4.4. Biosensor design and implementation in non-living protocells

In collaboration with Franck Molina (CNRS, Sys2Diag, Montpellier) and Jie-Hong Jiang (NTU, Taiwan) we ambition to apply our techniques to the design and implementation of biosensors in non-living vesicles for medical applications. Our approach is based on purely protein computation and on our ability to compile controllers and programs in biochemical reactions. The realization will be prototyped using a microfluidic device at CNRS Sys2Diag which will allow us to precisely control the size of the vesicles and the concentrations of the injected proteins. It is worth noting that the choice of non-living chassis, in contrast to living cells in synthetic biology, is particularly appealing for security considerations and compliance to forthcoming EU regulation.

5. Highlights of the Year

5.1. Highlights of the Year

- **Virtual Reality for Bacteria**
  Individual bacteria have been interfaced with a computer to build hybrid bio-digital circuits. Study published in Nature Communications [1].

- **Dynamical stabilization: real-time control allows maintaining cells in unstable configurations.**
  Using real-time control or periodic forcing one can drive cells towards a region of instability and dynamically maintain them there. Study published in Nature Communications [2].

- **Strong Turing Completeness of Continuous CRNs** solving a long standing open problem in CRN theory [8].

5.1.1. Awards

**BEST PAPER AWARD:**


6. New Software and Platforms

6.1. BIOCHAM

The Biochemical Abstract Machine

**KEYWORDS**: Systems Biology - Bioinformatics

**FUNCTIONAL DESCRIPTION**: The Biochemical Abstract Machine (BIOCHAM) is a software environment for modeling, analyzing and synthesizing biochemical reaction networks (CRNs) with respect to a formal specification of the observed or desired behavior of a biochemical system. BIOCHAM is compatible with the Systems Biology Markup Language (SBML) and contains some unique features about formal specifications in quantitative temporal logic, sensitivity and robustness analyses and parameter search in high dimension w.r.t. behavioral specifications, static analyses, and synthesis of CRNs.

**RELEASE FUNCTIONAL DESCRIPTION**: influence networks with forces – PAC learning of influence networks from time series data – synthesis of continuous reaction networks for mathematical functions defined by polynomial differential equations – complete modular rewriting of Biocham in SWI-Prolog

- Participants: François Fages, David Coudrin, Sylvain Soliman and Thierry Martinez
- Contact: François Fages
- URL: [http://lifeware.inria.fr/biocham/](http://lifeware.inria.fr/biocham/)

7. New Results

7.1. Strong Turing completeness of continuous CRNs

**Participants**: François Fages, Guillaume Le Guludec (former Member), Sylvain Soliman.

When seeking to understand how computation is carried out in the cell to maintain itself in its environment, process signals and make decisions, the continuous nature of protein interaction processes forces us to consider also analog computation models and mixed analog-digital computation programs. However, recent results in the theory of analog computability and complexity obtained by Pouly and Bournez, establish fundamental links between analog and digital computing. In [8] and [10], we derive from these results the strong (uniform computability) Turing completeness of chemical reaction networks over a finite set of molecular species under the differential semantics, solving a long standing open problem. Furthermore we derive from the proof a compiler of mathematical functions into elementary chemical reactions. We illustrate the reaction code generated by our compiler on trigonometric functions, and on various sigmoid functions which can serve as markers of presence or absence for implementing program control instructions in the cell and imperative programs. This makes it possible to start comparing our compiler-generated circuits to the natural circuit of the MAPK signaling network, which plays the role of an analog-digital converter in the cell with a Hill type sigmoid input/output functions.

7.2. Influence networks compared with reaction networks

**Participants**: François Fages, Thierry Martinez (former Member), David Rosenblueth (former Member), Sylvain Soliman, Denis Thieffry.
Biochemical reaction networks are one of the most widely used formalisms in systems biology to describe the molecular mechanisms of high-level cell processes. However, modellers also reason with influence diagrams to represent the positive and negative influences between molecular species and may find an influence network useful in the process of building a reaction network. In [11], we introduce a formalism of influence networks with forces, and equip it with a hierarchy of Boolean, Petri net, stochastic, and differential semantics, similarly to reaction networks with rates. We show that the expressive power of influence networks is the same as that of reaction networks under the differential semantics, but weaker under the discrete semantics. Furthermore, the hierarchy of semantics leads us to consider a (positive) Boolean semantics without test for absence, that we compare with the (negative) Boolean semantics with test for absence of gene regulatory networks à la Thomas. We study the monotonicity properties of the positive semantics and derive from them an algorithm to compute attractors in both the positive and negative Boolean semantics. We illustrate our results on models of the literature about the p53/Mdm2 DNA damage repair system, the circadian clock, and the influence of MAPK signaling on cell-fate decision in urinary bladder cancer.

7.3. Machine learning influence networks from data

Participants: Arthur Carcano, François Fages, Jérémy Grignard, Sylvain Soliman.

Automating the process of model building from experimental data is a very desirable goal to palliate the lack of modellers for many applications. However, despite the spectacular progress of machine learning techniques in data analytics, classification, clustering, and prediction making, learning dynamical models from data time-series is still challenging. In [7], we investigate the use of the Probably Approximately Correct (PAC) learning framework of Leslie Valiant as a method for the automated discovery of influence models of biochemical processes from Boolean and stochastic traces. We show that Thomas’ Boolean influence systems can be naturally represented by k-CNF formulae, and learned from time-series data with a number of Boolean activation samples per species quasi-linear in the precision of the learned model, and that positive Boolean influence systems can be represented by monotone DNF formulae and learned actively with both activation samples and oracle calls. We consider Boolean traces and Boolean abstractions of stochastic simulation traces, and show the space-time tradeoff there is between the diversity of initial states and the length of the time horizon, together with its impact on the error bounds provided by the PAC learning algorithms. We evaluate the performance of this approach on a model of T-lymphocyte differentiation, with and without prior knowledge, and discuss its merits as well as its limitations with respect to realistic experiments.

7.4. Shaping bacterial population behavior through computer-interfaced control of individual cells

Participant: Jakob Ruess.

Bacteria in groups vary individually, and interact with other bacteria and the environment to produce population-level patterns of gene expression. Investigating such behavior in detail requires measuring and controlling populations at the single-cell level alongside precisely specified interactions and environmental characteristics. In [1], we present an automated, programmable platform that combines image-based gene expression and growth measurements with on-line optogenetic expression control for hundreds of individual Escherichia coli cells over days, in a dynamically adjustable environment. This integrated platform broadly enables experiments that bridge individual and population behaviors. We demonstrate: (i) population structuring by independent closed-loop control of gene expression in many individual cells, (ii) cell–cell variation control during antibiotic perturbation, (iii) hybrid bio-digital circuits in single cells, and freely specifiable digital communication between individual bacteria. These examples showcase the potential for real-time integration of theoretical models with measurement and control of many individual cells to investigate and engineer microbial population behavior.

7.5. Balancing a genetic toggle switch by real-time feedback control and periodic forcing

Participants: Gregory Batt, Jean-Baptiste Lugagne, Melanie Kirch (former Member), Agnes Köhler (former Member), Sebastian Sosa Carrillo.
Cybergenetics is a novel field of research aiming at remotely pilot cellular processes in real-time with to leverage the biotechnological potential of synthetic biology. Yet, the control of only a small number of genetic circuits has been tested so far. Here we investigate the control of multistable gene regulatory networks, which are ubiquitously found in nature and play critical roles in cell differentiation and decision-making. Using an in silico feedback control loop, we demonstrate that a bistable genetic toggle switch can be dynamically maintained near its unstable equilibrium position for extended periods of time [2]. Importantly, we show that a direct method based on dual periodic forcing is sufficient to simultaneously maintain many cells in this undecided state. These findings pave the way for the control of more complex cell decision-making systems at both the single cell and the population levels, with vast fundamental and biotechnological applications.

7.6. Abstracting the dynamics of biological pathways using information theory: a case study of apoptosis pathway

*Participants:* Gregory Batt, François Bertaux (former Member), Sucheendra Palaniappan (former Member).

Quantitative models are increasingly used in systems biology. Usually, these quantitative models involve many molecular species and their associated reactions. When simulating a tissue with thousands of cells, using these large models becomes computationally and time limiting. In our paper, we propose to construct abstractions using information theory notions [3]. Entropy is used to discretize the state space and mutual information is used to select a subset of all original variables and their mutual dependencies. We apply our method to an hybrid model of TRAIL-induced apoptosis in HeLa cell. Our abstraction, represented as a Dynamic Bayesian Network (DBN), reduces the number of variables from 92 to 10, and accelerates numerical simulation by an order of magnitude, yet preserving essential features of cell death time distributions.

7.7. Long-term tracking of budding yeast cells in brightfield microscopy: CellStar and the Evaluation Platform

*Participants:* Gregory Batt, Artémis Llamosi (former Member).

With the continuous expansion of single cell biology, the observation of the behaviour of individual cells over extended durations and with high accuracy has become a problem of central importance. Surprisingly, even for yeast cells that have relatively regular shapes, no solution has been proposed that reaches the high quality required for long-term experiments for segmentation and tracking (S&T) based on brightfield images. In this contribution, we present CellStar, a tool chain designed to achieve good performance in long-term experiments [5]. The key features are the use of a new variant of parametrized active rays for segmentation, a neighbourhood-preserving criterion for tracking, and the use of an iterative approach that incrementally improves S&T quality. A graphical user interface enables manual corrections of S&T errors and their use for the automated correction of other, related errors and for parameter learning. We created a benchmark dataset with manually analysed images and compared CellStar with six other tools, showing its high performance, notably in long-term tracking. As a community effort, we set up a website, the Yeast Image Toolkit, with the benchmark and the Evaluation Platform to gather this and additional information provided by others.

7.8. Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability

*Participant:* Jakob Ruess.

Determining the sensitivity of certain system states or outputs to variations in parameters facilitates our understanding of the inner working of that system and is an essential design tool for the de novo construction of robust systems. In cell biology, the output of interest is often the response of a certain reaction network to some input (e.g., stressors or nutrients) and one aims to quantify the sensitivity of this response in the presence of parameter heterogeneity. We argue that for such applications, parametric sensitivities in their standard form do not paint a complete picture of a system’s robustness since one assumes that all cells in the population have the same parameters and are perturbed in the same way. In the published contribution, we consider stochastic
reaction networks in which the parameters are randomly distributed over the population and propose a new sensitivity index that captures the robustness of system outputs upon changes in the characteristics of the parameter distribution, rather than the parameters themselves [4]. Subsequently, we make use of Girsanov’s likelihood ratio method to construct a Monte Carlo estimator of this sensitivity index. However, it turns out that this estimator has an exceedingly large variance. To overcome this problem, we propose a novel estimation algorithm that makes use of a marginalization of the path distribution of stochastic reaction networks and leads to Rao-Blackwellized estimators with reduced variance.

7.9. Recombinase-based genetic circuit optimization by integer linear programming

Participant: François Fages.

The rapid advancements of synthetic biology show promising potential in biomedical and other applications. Recently, recombinases were proposed as a tool to engineer genetic logic circuits with long-term memory in living and even mammalian cells. The technology is under active development, and the complexity of engineered genetic circuits grows continuously. However, how to minimize a genetic circuit composed of recombinase-based logic gates remain largely open. In [12], we formulate the problem as a cubic-time assignment problem and solved by a 0/1-ILP solver to minimize DNA sequence length of genetic circuits. Experimental results show effective reduction of our optimization method, which may be crucial to enable practical realization of complex genetic circuits.

7.10. Coupled models of the cell cycle and circadian clock

Participants: François Fages, Sylvain Soliman, Pauline Traynard (former Member).

Experimental observations have put in evidence autonomous self-sustained circadian oscillators in most mammalian cells, and proved the existence of molecular links between the circadian clock and the cell cycle. Some mathematical models have also been built to assess conditions of control of the cell cycle by the circadian clock, with applications to cancer chronotherapy optimization. However, recent studies in individual NIH3T3 fibroblasts have shown an unexpected acceleration of the circadian clock together with the cell cycle when the culture medium is enriched with growth factors, and the absence of such acceleration in confluent cells. In order to explain these observations, we have studied a possible entrainment of the circadian clock by the cell cycle through a regulation of clock genes around the mitosis phase. We developed a computational model in Biocham with a formal specification of the observed behavior in quantitative temporal logic to investigate the conditions of entrainment in period and phase. We showed that either the selective activation of RevErb-α or the selective inhibition of Bmal1 transcription during the mitosis phase, allowed us to fit the experimental data on both period and phase, while a uniform inhibition of transcription during mitosis seems incompatible with the phase data. In [6], we presented those results and some further predictions of the bidirectional model with a coupling in both directions.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR Projects

- ANR-MOST BIOPSY (2016-2020) on “Biochemical Programming System”, coordinated by F. Molina (CNRS, Sys2diag, Montpellier) and J.H. Jiang (National Taiwan University), with F. Fages.
- ANR MEMIP (2016-2020) on “Mixed-Effects Models of Intracellular Processes”, coordinated by G. Batt, with P. Hersen, (CNRS/Paris7), E. Cinquemani (Inria EPI IBIS) and M. Lavielle (Inria/CNRS/Polytechnique, EPI XPOP).
• ANR COGEX (2016-2019) on “Computer Aided Control of Gene Expression” coordinated by P. Hersen (MSC lab, CNRS/Paris7), with G. Batt and G. Truan (LISBP, CNRS/INSa).

• ANR Blanc HYCLOCK (2014-2018) on “Hybrid modeling of time for Circadian Clock Biology and Chronopharmacology”, coordinated by F. Delaunay (CNRS, Nice), with F. Lévi (INSERM Paris-Sud), G. Bernot (CNRS I3S, Nice), O. Roux (Ecole Centrale Nantes), F. Fages and S. Soliman.

• ANR Blanc STOCH-MC (2014-2018) on “Stochastic Models: Scalable Model Checking”, coordinated by Blaise Genest (Inria Rennes), with Grégory Batt, Wiesław Zielonka (LIAFA), and Hugo Gimbert (LabRI).

• ANR Investissement Avenir ICEBERG project (2011-2017) “From population models to model populations”, coordinated by Grégory Batt, with Pascal Hersen (MSC lab, Paris Diderot Univ./CNRS), Reiner Veitia (Institut Jacques Monod, Paris Diderot Univ./CNRS), Olivier Gandrillon (BM2A lab, Lyon Univ./CNRS), Cédric Lhoussaine (LIFL/CNRS), and Jean Krivine (PPS lab, Paris Diderot Univ./CNRS).

8.1.2. Inria Project Lab

• IPL COSY (2017-2021) “real-time control of synthetic microbial communities”, coordinated by Eugenio Cinquemani (Ibis, Inria), with Jean-Luc Gouzé (Biocoore, Inria), Gregory Batt, Frédéric Bonnans (Commands, Inria), Efimov Denis (Non-A, Inria), and Hans Geiselmann (BIOP, Université Grenoble-Alpes), Beatrice Laroche (Maiage, Inra Jouy-en-Josas), and Hyun Youk (Youk lab, TU Delft).

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

• H2020 FET-OPEN COSY-BIO (2017-2020), “Control Engineering of Biological Systems for Reliable Synthetic Biology Applications”, coordinated by Diego di Bernardo (Tigem), with Filippo Menolascina (Edinburgh U), Mario di Bernardo (Naples U), Pascal Hersen (Paris7 U), Mustafa Khammash (ETHZ), Gregory Batt, Guy-Bart Stan (Imperial College), and Lucia Marucci (Bristol U).

8.3. International Initiatives

8.3.1. Participation in International Programs


9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

François Fages is co-organizer of the CSBC Workshop on Computational Systems Biology for Cancer, Institut des Systèmes Complexes, Paris, France, 24-26 jan. 2018.
9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- François Fages was member of the PC of
  - **CMSB'17** The 15th conference on Computational Methods for Systems Biology, September 27th to 29th 2017, Darmstadt, Germany.
  - **FroCos’17** The 11th International Symposium on Frontiers of Combining Systems, 25-29 September 2017, Brasilia, Brazil.

9.1.2.2. Reviewer

- François Fages reviewed one article for the 15th Int. Computer Science Symposium in Russia CSR 2017.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

François Fages is member of

- the Editorial Board of the Computer Science area of the Royal Society Open Science journal since 2014,
- the Editorial Board of the journal RAIRO OR Operations Research since 2004.

9.1.3.2. Reviewer - Reviewing Activities

In addition to their Editorial Board and Program Committee duties,

- Grégory Batt reviewed one article for *Nature Scientific Reports*.
- François Fages reviewed articles for *Bioinformatics, ACM Transactions on Modeling and Computer Simulation, Fundamenta Informaticae, Computers and Industrial Engineering*
- Sylvain Soliman reviewed articles for *Briefings in Bioinformatics, PNAS, Transactions on Computational Biology and Bioinformatics*

9.1.4. Invited Talks

- Virgile Andréani gave a poster presentation at the Spring school on Computational Systems Biology (CompSysBio 2017), Aussois, March 19-25 2017
- Gregory Batt gave invited talks at
  - Workshop on Control of Cellular and Molecular Systems, “Balancing a genetic toggle switch by real-time control or periodic stimulations”, Mathematical Biosciences Institute, Ohio State University.
François Fages gave invited talks at

- GT BIOSS-preGDR IA, keynote talk “In Quest of the Software of the Living: Successes and Difficulties of the Program Verification Paradigm in Cell Biology”, Gif sur Yvette, 22 June 2017.
- “Turing Completeness of Biochemical Reactions over a finite set of molecules under the Differential Semantics and Compilation of Mixed Analog-Digital Programs”, French Embassy in Berlin, Germany, 1 June 2017.

Jean-Baptiste Lugagne gave invited talks at

- Khammash lab seminar, “Balancing a genetic toggle switch by real-time feedback control and periodic forcing”, ETHZ, Zurich, March 2017
- Seminar of SynthSys Centre for Synthetic and Systems Biology, “Balancing a genetic toggle switch by real-time feedback control and periodic forcing”, University of Edinburgh, June 08 2017
- International Workshop on Control Engineering and Synthetic Biology, ‘Balancing a genetic toggle switch by real-time feedback control and periodic forcing” (poster presentation), Royal Academy of Engineering, London, July 17 2017
- Seminar of Theory of Living Matter Group, “Balancing a genetic toggle switch by real-time feedback control and periodic forcing”, University of Cambridge, July 19 2017

Jakob Ruess gave invited talks at

- Working days of GT BIOSS “Control of bio-digital systems in single cells”, March 13 2017

9.1.5. Leadership within the Scientific Community

- Grégory Batt is a member of
  - the IEEE/CSS Technical Committee on Systems Biology,
  - the scientific board of the GDR de Biologie de Synthèse et des Systèmes
  - the GDR de Bioinformatique Moléculaire, in charge of the axis on Biological network modelling, systems biology and synthetic biology
  - co-animator of the French working group on Symbolic Systems Biology GT BIOSS
  - the scientific committee of the Spring school on Computational Systems Biology (Comp-SysBio 2017)

- François Fages is a member of
– the Steering Committee of the International Conference on Computational Methods for Systems Biology since 2008,
– the Scientific Council of the Doctorate School “Frontières Du Vivant” at Center for Research and Interdisciplanirity, Universities Paris Descartes and Paris Diderot, since 2010,

9.1.6. Scientific Expertise

• Gregory Batt
  – has been in charge of the definition of the 16th challenge of the 2018-2022 Inria strategic plan on predictive systems biology
  – has been a member of the INRA selection committee for hiring permanent junior researchers (CR2)
  – has evaluated the PhD thesis of Lorena Postiglioni and Giansimone Perrino (Diego di Bernardo, Tigem, Italy; Thesis evaluation report)
  – has been a member of thesis advisory committee meetings for Arnaud Bonnaffoux (Olivier Gandrillon ENS Lyon/Cosmo Company), Raphael Goujet (Ariel Lindner, CRI, Paris), Adele Kerjouan (Olivier Destaing, IAB, Grenoble), and Mathilde Koch (Jean-Loup Faulon, INRA, Jouy en Josas)

• François Fages
  – is a member of the jury for the Prix de thèse Gilles Kahn of the Société Informatique de France, since 2015,
  – reviewed one Research Fellowship application for the Royal Commission for the Exhibition of 1851

9.1.7. Research Administration

• François Fages is member of the “Bureau du Comité des Projets” of Inria Saclay-IdF.
• Sylvain Soliman is member of the “Commission Scientifique” of Inria Saclay-IdF

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Grégory Batt (coordinator and teacher: 48h) and Jakob Ruess (24h), Computational Biology, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master/PhD: Grégory Batt (co-coordinator 80h, teacher 8h) and Jakob Ruess (8h), Modeling and engineering of biological systems, M2/PhD, Institut de Technologie et d’Innovation of Paris Sciences et Lettres (PSL-ITI), Paris.

Master/PhD: Gregory Batt (Scientific committee and teacher 6h), Spring school on Computational Systems Biology (CompSysBio 2017), Aussois, March 19-25 2017

Master: François Fages (coordinator module 48h, teaching 12h), C2-19 Computational Methods for Systemic and Synthetic Biology, Master Parisien de Recherche en Informatique (MPRI), Paris.

Master: Chiara Fracassi (20h), Experimental Methods in Biophysics, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master: Sylvain Soliman, C2-35-1 Constraint Programming, coordinator and teaching 24h, M2, Master Parisien de Recherche en Informatique (MPRI), Paris.

Master: Denis Thieffry (coordinator, 6h) and Gregory Batt (6h), Dynamical Modelling of Cellular Regulatory Networks, M2, Interdisciplinary Master in Life Science at the Ecole Normale Supérieure, Paris.
E-learning
Online BIOCHAM tutorial notebook by F. Fages and S. Soliman, presented at CMSB 2017, Darmstadt, Germany.

9.2.2. Supervision
Grégory Batt is currently supervising the Ph.D. theses of
- Virgile Andreani (ED STIC, ENS)
- Sebastian Sosa Carrillo (ED FdV, Inria)
- Chetan Aditya (ED FdV, Inria)

9.2.3. Juries
- Ph.D.: Guillaume Madelaine, University of Lille, F. Fages, reviewer, Feb. 28 2017.
- Ph.D.: Danhua Peng, University of Rostock, Germany, F. Fages, reviewer, Feb. 3 2017.

9.3. Popularization
We participated in seminars for general audiences.
- J. Ruess, “Automating scientific discovery in single cells”, Unité ou café, Inria Saclay - Ile de France, Apr 21 2017

Our publications also raised the attention of specialized and non-specialized media. One can mention
- “Piloter le comportement de cellules biologiques par ordinateur, c’est possible” in Science et Avenir
- “Forscher verbinden und kontrollieren Bakterien mit einem Computer”, in Wiener Zeitung
- “Forscher steuern Bakterien durch Verbindung mit Computer” in der Standard
- “Forscher verbinden und kontrollieren Bakterien mit einem Computer” and “Virtuelle Realität für Bakterien” in Austria Presse Agentur (APA)
- “Ils parviennent à contrôler des cellules biologiques par ordinateur” in SciencePost
- “Computerized biology, or how to control a population of cells with a computer” and “Virtual reality for bacteria” in Phys.org
- “Computerized biology, or how to control a population of cells with a computer” in EurekAlert.org
- “Scientists Learn to Control Bacteria by Connecting Them to Computers” in Scicasts.com
- “Virtuelle Realität für Bakterien” in Innovations-report.de
- “Virtuelle Realität für Bakterien” in Informationsdienst Wissenschaft

10. Bibliography

Publications of the year
Articles in International Peer-Reviewed Journals


Invited Conferences


International Conferences with Proceedings


[8] Best Paper

Scientific Books (or Scientific Book chapters)


Scientific Popularization

Other Publications
