

Modélisation et analyse par vérification formelle et simulation conjointe des réseaux de régulation géniques

Olivier Roux

IRCCyN, (UMR CNRS 6597)

1 rue de la Noë, BP 92101 F-44321 Nantes Cedex 3 (FRANCE)

roux@irccyn.ec-nantes.fr

Résumé Pour comprendre les évolutions dans les systèmes vivants, les contrôler et éventuellement éviter ou remédier à certains comportements mauvais, la biologie synthétique s'attache à modéliser ces processus dynamiques afin de les analyser précisément. Plusieurs types de modélisation sont mis en œuvre, parmi lesquels notre démarche se caractérise par une *approche hybride* (mélangeant des aspects discrets et continus) basée sur la formalisation discrète de René Thomas, et elle s'exprime en utilisant l'approche algébrique et temporelle du *π -calcul stochastique*. Nous illustrons succinctement les résultats obtenus sur un exemple simple et nous montrons les bénéfices que nous pouvons tirer de cette nouvelle formulation.

Abstract. In order to understand development of living organisms, and to control them and possibly to avoid or remedy some unexpected behaviours, synthetic biology intends to build models of these dynamical processes to accurately analyse them. Several modelling frameworks were designed for such a purpose, among which the characteristic of our approach is that it leads to *hybrid models* (mixing discrete and continuous features) which are founded upon the discrete formalism of René Thomas. Furthermore, it is algebraically expressed thanks to the *stochastic π -calculus* including the temporal aspects. We shortly highlight the results achieved from a simple example and we show the benefits that we brought from this new method.

1 Introduction

Que ce soit au niveau intra-cellulaire, ou à d'autres échelles, les phénomènes biologiques révèlent des comportements très variables d'entités dynamiques interagissant de façon complexe. Pour comprendre les évolutions des systèmes vivants, les contrôler et éventuellement éviter ou remédier à certains comportements mauvais, la biologie synthétique s'attache à modéliser ces processus dynamiques afin de les analyser précisément. Ces analyses consistent à imaginer et mettre en œuvre les prévisions d'évolutions du système en fonction de nombreux paramètres de natures très variées, à la fois internes au système considéré et révélateurs de l'environnement de ce système. La complexité vient notamment de la diversité des actions/réactions mutuelles des entités et de leur multiplicité, ainsi que du très grand nombre de paramètres dont la connaissance des valeurs est difficile à obtenir mais essentielle pour mener à bien les études envisagées.

Plusieurs démarches de modélisation sont mises en application, parmi lesquelles les équations différentielles constituent une approche classique dans ce type de recherche, mais qui se heurte, de façon plus ou moins rédhibitoire, au délicat problème

d'estimation d'un nombre important de paramètres. Une autre approche, que nous privilégions, consiste à s'attacher à une représentation par des systèmes états-transitions. Ce mécanisme se prête assez bien à des extensions parmi lesquelles celle qui prend en compte la dimension temporelle des phénomènes — non seulement par les aspects qualitatifs (chronologiques), mais aussi quantitatifs (chronométriques) — nous paraît particulièrement adaptée.

Sans en constituer une synthèse, les idées prospectives, originales et pas encore publiées, que nous présentons ici se basent sur une somme importante d'expériences déjà menées ou en cours sur la modélisation et la vérification de ces réseaux génétiques et avec passage à l'échelle.

Nous présentons notre démarche dans la section suivante en décrivant le cadre de modélisation que nous avons établi sur une *approche hybride* (mélangeant des aspects discrets et continus) et par une expression en π -calcul stochastique [7]. Puis, nous montrons les résultats de simulation que nous avons obtenus sur un exemple très simple (à partir d'une modélisation à six états de la régulation de deux gènes intervenant de façon déterminante dans les comportements de la bactérie pathogène *Pseudomonas aeruginosa*). Avant de conclure, nous indiquons comment tirer profit de cette nouvelle représentation algébrique.

2 Description du modèle

Fondamentalement, notre démarche, qui se place dans la lignée du principe de modélisation discrète de René Thomas [9], renforce les travaux que nous avons effectués et proposés [1,2] sur la modélisation hybride. Comme nous allons le montrer, ce récent travail et les perspectives qu'il ouvre permettent d'affiner certains résultats que nous obtenions dans la première phase de notre démarche. Il se place en outre dans l'esprit des recherches menées actuellement dans notre équipe [3,4,5].

2.1 Motivations

Nous partons donc d'un espace d'états où chaque état est constitué, selon le principe de René Thomas, à partir de l'ensemble des valeurs abstraites qui concernent, pour chaque gène, les seuils (discrets) au delà desquels il commence à exercer une influence positive (activateur) ou négative (inhibiteur) sur un autre gène. Notre originalité consiste à *considérer le temps qui s'écoule dans chacun de ces états* jusqu'au passage dans un autre état au bout d'une durée donnée par des paramètres de délais de l'activation ou de l'inhibition de chaque gène. Dans la modélisation hybride que nous avons ainsi conçue et présentée dans [1,2], le nombre d'états et de paramètres mis en jeu est très important et ceci constitue actuellement un facteur limitatif très contraignant pour la plupart des analyses que nous voudrions pouvoir effectuer sur des applications biologiques réelles. C'est pourquoi nous avons imaginé la traduction dont nous allons présenter ici les principes et qui nous permet, d'une part, d'effectuer des simulations pouvant induire certains résultats et, d'autre part de déboucher, grâce à des transformations algébriques, sur des simplifications intéressantes.

2.2 Principe de la traduction en π -calcul

Le principe de traduction est assez simple et direct puisqu'il se résume à transformer chaque état en un processus du π -calcul stochastique [8,7] qui s'exécute en passant un certain délai avant de se transformer en un autre processus correspondant à un nouvel état vers lequel il était susceptible de transiter. De fait, la transition effectivement choisie dépend des délais de transitions qui se trouvent donc être des paramètres comme dans la modélisation hybride. Pourtant, nous en tirons profit des façons suivantes.

2.3 Résultats de simulation

L'application de la démarche esquissée ci-dessus, en utilisant l'outil SPiM [6], donne des résultats de simulation (ici, par exemple, sur cent exécutions) qui sont illustrés par la figure 1. Deux états particuliers du système ont été sélectionnés : d'une part, un état, choisi comme état initial, et appartenant à un cycle potentiel d'exécution ; et d'autre part, un état puits (à partir duquel il n'y a pas de transition vers un autre état). En fonction du temps, figure le nombre d'exécutions se trouvant dans l'un des deux états considérés¹. Pour rendre les résultats plus lisibles, nous avons exclu les quatre autres états du modèle discret.

Ces résultats correspondent à des valeurs de paramètres de délai fixés plus ou moins arbitrairement mais qui mettent en évidence deux types de comportements fondamentalement différents, c'est-à-dire conduisant dans un cas à une attraction très forte vers un état d'absorption et la diminution du nombre de passages par l'état appartenant au cycle (à gauche), et un équilibre entre les exécutions conduisant vers cet état d'absorption ou à des répétitions sempiternelles de passage par le même état du cycle (à droite).

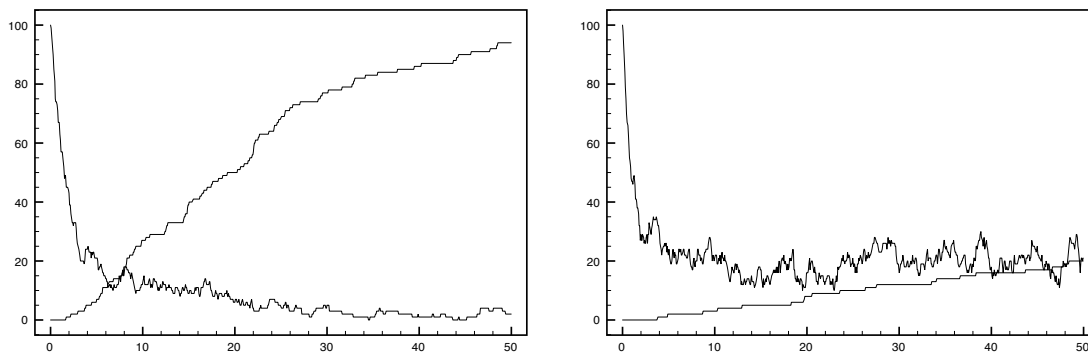


Figure 1 : Résultats obtenus pour différents délais d'activation et inhibition

2.4 Simplifications algébriques

Un avantage de la formulation algébrique (en π -calcul) du comportement global est aussi qu'elle permet de mettre en œuvre des opérateurs pour calculer des équivalences.

¹ A l'instant initial, toutes les exécutions sont donc dans l'état initial.

En particulier, nous pensons que les cycles, qui correspondent à des phénomènes oscillatoires, pourraient être reconnus comme tels et, dans certains cas, être abstraits en des phases dans lesquels on ne s'intéresse qu'au temps qui s'écoule dans ces phases. De même, alors qu'on ne considère actuellement qu'un unique processus, résultant de la mise à plat de l'ensemble des exécutions, on pourrait faire en sorte d'ajouter des mises en parallèle avec des comportements externes de type environnemental ou multi-échelle.

3 Conclusion

Les travaux brièvement présentés ici sont évidemment loin d'être définitivement aboutis, mais ils constituent, au contraire, un état provisoire de recherches en cours et ils se placent naturellement dans une démarche cohérente allant de la modélisation des systèmes complexes de régulation d'après René Thomas à la prise en compte des délais d'interaction jusqu'à une représentation algébrique permettant à la fois des vérifications formelles et des simulations pour la synthèse de paramètres. Précisément, cette cohérence de diverses approches laisse entrevoir des perspectives assez riches, notamment pour arriver à pallier les problèmes d'explosion de l'espace d'analyse, qui sont dûs à la complexité des systèmes considérés.

Références

1. Jamil Ahmad, Gilles Bernot, Jean-Paul Comet, Didier Lime, and Olivier Roux. Hybrid modeling and dynamical analysis of gene regulatory networks with delays. *ComplexUs*, 3(4) :231–251, October 2007.
2. Jamil Ahmad and Olivier Roux. Invariance kernel of biological regulatory networks. In *IEEE Int. Conference on Bioinformatics and Biomedicine (BIBM 2008)*, pages 135–140, Philadelphia (Penn.), USA, nov 2008. IEEE Computer Society.
3. Jonathan Fromentin, Jean-Paul Comet, Pascale Le Gall, and Olivier Roux. Analysing gene regulatory networks by both constraint programming and model-checking. In *EMBC07, 29th IEEE EMBS Annual International Conference*, pages 4595–4598, August 23–26 2007.
4. Mylène Maurin, Morgan Magnin, and Olivier Roux. Modeling of genetic regulatory network in stochastic pi-calculus. In *International Conference on Bioinformatics and Computational Biology (BICOB 2009)*, LNBI, New Orleans (Louisiana) USA, apr 2009. Springer Verlag. to appear.
5. Loïc Paulevé, Morgan Magnin, and Olivier Roux. Refining dynamics of gene regulatory networks in a stochastic pi-calculus framework, 2009. submitted.
6. Andrew Phillips. SPiM, 2007. <http://research.microsoft.com/~aphillip/spim>.
7. Andrew Phillips and Luca Cardelli. Efficient, correct simulation of biological processes in the stochastic pi-calculus. In *Computational Methods in Systems Biology (CMSB 2007)*, volume 4695 of *LNCS*, pages 184–199. Springer, September 2007.
8. Corrado Priami. Stochastic pi-calculus. *The Computer Journal*, 38(7) :578–589, 1995.
9. René Thomas, Denis Thieffry, and Marceline Kaufman. Dynamical behaviour of biological regulatory networks—i. biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull Math Biol*, 57(2) :247–276, Mar 1995.