

# Satisfaction de problèmes logiques hybrides et application à la synthèse de règles de régulation métabolique

Kerian Thuillier<sup>1</sup>, Loïc Paulevé<sup>2</sup>, Anne Siegel<sup>1</sup>

<sup>1</sup> Univ Rennes, Inria, CNRS, IRISA, F-35000 Rennes, France  
`{prénom}.{nom}@irisa.fr`

<sup>2</sup> Univ. Bordeaux, Bordeaux INP, CNRS, LaBRI, UMR5800, F-33400 Talence, France  
`{prénom}.{nom}@labri.fr`

**Mots-clés :** *Inférence de régulation, Satisfiabilité Modulo Théorie, Biologie des systèmes*

## 1 Introduction

La régulation de l'expression des gènes métaboliques est essentielle à l'adaptation de la cellule à son environnement. Depuis 30 ans, différentes méthodes de modélisation, simulation et inférences des réseaux de régulation de gènes ont été proposées. Cependant, ces réseaux restent, à ce jour, très largement incomplets ne permettant pas la prédiction précise des réponses des cellules face à un changement d'environnement.

Une cellule est couramment représentée par deux systèmes inter-connectés : un système métabolique, un ensemble de réactions chimiques et métabolites, et un système de régulation et signalisation, un ensemble de règles permettant à la cellule de s'adapter à son environnement. Malgré l'inter-connexion des deux systèmes, l'inférence des réseaux de régulation se fait classiquement en ne considérant que le système de régulation et signalisation. Les quelques approches d'inférence basées sur les deux systèmes ne considèrent, quant à elles, que les interactions du système de régulation sur le système métabolique. Elles manquent certaines régulations dépendant du retour du métabolisme sur le système de régulation.

Dans ces travaux, nous présentons une nouvelle approche pour inférer les règles booléennes de régulation du métabolisme à partir de séries temporelles et de connaissances *a priori*, modélisées par un réseau de connaissance *a priori* (*Prior Knowledge Network* — PKN). Notre méthode est basée sur la combinaison d'un solveur *Answer Set Programming* (ASP) et d'un solveur linéaire. En résolvant à la fois des contraintes combinatoires et d'arithmétique linéaire, nous générons des règles booléennes de régulation admissible vis-à-vis des séries temporelles et du PKN d'entrée.

## 2 Inférence de réseaux de régulation

**Réseaux métaboliques régulés.** Les cellules sont généralement modélisées par des réseaux métaboliques régulés (RMN). Un *réseau métabolique régulé* (RMN) est une paire  $N, f$  où :

1.  $N$  est un réseau métabolique avec  $N = (R \cup M, s)$  où  $R$  est l'ensemble des réactions,  $M$  l'ensemble des métabolites et  $s : R \times M \rightarrow \mathbb{R}$  la fonction de stoechiométrie. Le réseau métabolique est caractérisé par des contraintes linéaires sur les flux métaboliques.
2.  $f$  est un réseau booléen modélisant le système de régulation avec  $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$  où  $n$  le nombre d'éléments du réseau de régulation.

**Dynamique.** Ces deux systèmes ont des dynamiques et des échelles de temps très différentes. Le système métabolique est un système rapide caractérisé par des contraintes linéaires. Sa simulation repose sur la résolution du problème de programmation linéaire Eq .1 :

$$\begin{aligned}
& \text{maximiser} && v_{\text{obj}} \\
& \text{sous contraintes :} && \\
& \forall m \in M, && \sum_{r \in R} s(r, m) \times v_r = 0 \\
& \forall r \in R, && x_r \times l_r \leq v_r \leq u_r \times x_r
\end{aligned} \tag{1}$$

où  $\forall r \in R, v_r \in \mathcal{R}$  est une variable linéaire modélisant le flux métabolique de la réaction  $r$ ,  $l_r$  et  $u_r$  sont, respectivement, les bornes inférieure et supérieure de son flux métabolique,  $x_r \in \{0, 1\}$  est son état (inhibée ou non), et  $\text{obj} \in R$  est une fonction de biomasse. Tandis que le système de régulation est un système lent suivant une dynamique discrète. Le nouvel état booléen  $y \in \{0, 1\}^n$  est calculé à partir de l'état précédent  $x \in \{0, 1\}^n$  tel que  $y = f(x)$ .

La dynamique de la cellule est hybride, *i.e.* discrète et linéaire, et peut être simulée, avec des approches comme la rFBA [2] en alternant les mises-à-jour successives des deux systèmes.

**Problème d'inférence de régulation métabolique.** Le problème d'inférence des règles de régulation métabolique est défini comme :

- Entrées :** un réseau métabolique  $N = (R \cup M, s)$ , un domaine de recherche  $\mathbb{D}$  défini par un ensemble de contraintes combinatoires et une séquence de  $k$  observations  $\{(x^t, v_{\text{obj}}^t)\}_{t=0}^k$ .
- Sorties :** l'ensemble des réseaux booléens  $f \in \mathbb{D}$  tel qu'il existe une simulation rFBA du réseau métabolique régulé  $(N, f)$  consistante avec la séquence des  $k$  observations.

### 3 Méthode de résolution

**Méthode de résolution.** Pour résoudre le problème d'inférence, nous proposons une approche [3] basée sur les principes des solveurs *Satisfiability Modulo Théorie* (SMT) et implémentant un système de propagation de contraintes [1]. Le problème d'inférence est divisé en deux sous-problèmes : l'un purement combinatoire qui est résolu par un solveur ASP, et l'autre un problème linéaire qui est résolu par un solveur linéaire. Le premier problème va générer des réseaux de régulation candidats. Pour chaque candidat, on construit les problèmes linéaires associés ; *i.e.* Eq. 1 avec l'état  $x \in \{0, 1\}^{|R|}$  des réactions fixé. Ils sont résolus à l'aide d'un solveur linéaire. Si pour chaque problème linéaire, l'optimum du problème respecte l'observation en entrée associée, le candidat est accepté. Sinon, il est rejeté et de nouvelles contraintes combinatoires sont générées et ajoutées au solveur ASP.

**Propagation de contraintes.** Ces nouvelles contraintes sont basées sur des propriétés monotones liant l'ensemble des contraintes d'un problème linéaire à une borne supérieure et/ou inférieure sur son optimum. La propriété dit que : étant donné un problème linéaire dont on connaît l'optimum, retirer (resp. ajouter) une contrainte linéaire ne peut pas diminuer (resp. augmenter) l'optimum. Les contraintes combinatoires ainsi générées permettent d'empêcher le solveur ASP de générer des réseaux booléens candidats qui échoueront la vérification linéaire.

### Références

- [1] Clark Barrett and Cesare Tinelli. *Satisfiability Modulo Theories*, pages 305–343. Springer International Publishing, Cham, 2018.
- [2] Lucas Marmiesse, Rémi Peyraud, and Ludovic Cottret. FlexFlux : combining metabolic flux and regulatory network analyses. *BMC Systems Biology*, 9(1), 2015.
- [3] Kerian Thuillier, Caroline Baroukh, Alexander Bockmayr, Ludovic Cottret, Loïc Paulevé, and Anne Siegel. MERRIN : METabolic regulation rule INFerence from time series data. *Bioinformatics*, 38(Supplement\_2) :ii127–ii133, 09 2022.