

# Formalisation des réseaux biomoléculaires complexes

Ali Ayadi\*, François de Beuvron\* Cecilia Zanni-Merk\*, Julie Thompson\*

\*ICube (UMR CNRS 7357) - 300 bd Sébastien Brant - BP 10413 - F-67412 Illkirch Cedex  
{ali.ayadi,debeuvron,merk,thompson}@unistra.fr,

**Résumé.** La "transitabilité" d'un réseau biomoléculaire complexe exprime l'idée du pilotage de ce réseau d'un état non souhaité (en général associé à une maladie) à un état désiré. Dans ce article, nous présentons une formalisation du réseau biomoléculaire complexe et nous proposons une plate-forme basée sur les technologies sémantiques pour l'optimisation de la transitabilité de ces réseaux biomoléculaires.

## 1 Introduction

Dans le domaine de la biologie moléculaire, la conception et l'exploitation de nouvelles approches expérimentales s'intéressant au suivi de la structure des composants cellulaires et à leur construction ouvrent la voie à un nouveau domaine d'étude de la bioinformatique, la "Transitabilité". Selon [Wu et al. \(2014\)](#), *la transitabilité est l'orientation ou le pilotage d'un réseau biomoléculaire complexe d'un état non satisfaisant à un état désiré.*

Un réseau biomoléculaire est orchestré par les interactions de plusieurs molécules dans une cellule. Par défaut, une cellule vivante doit rester à un état normal (phénotype sain). Cependant, par une certaine perturbation inconnue ou stimulus, le réseau biomoléculaire peut changer d'un phénotype sain à un phénotype malade. Il est ainsi indispensable de trouver des mécanismes pour faire évoluer le réseau biomoléculaire du phénotype anormal au phénotype sain. Pour étudier les transitions de phénotypes, le réseau biomoléculaire est modélisé par un graphe dans lequel les molécules sont représentés par des nœuds et les interactions entre les molécules sont représentées par des arcs. En conséquence, les phénotypes cellulaires peuvent être définis par les états du réseau. Un état du réseau représente les expressions de toutes les molécules intervenant dans le réseau. Les changements phénotypiques ou les changements de comportement cellulaire seront donc décrits par une transition dynamique entre deux états du réseau [Wu et al. \(2014\)](#).

C'est dans ce contexte que nos travaux de thèse s'articulent autour du développement d'une plate-forme d'optimisation de la transitabilité des réseaux biomoléculaires complexes en proposant des mécanismes de pilotage des transitions de ces réseaux d'un état quelconque à un état spécifique. Nos travaux ont débuté en mai 2015, ayant comme objectif d'explorer le potentiel des techniques d'optimisation combinatoire et sémantiques pour optimiser et contrôler un réseau biomoléculaire complexe.

## 2 Modélisation du réseau biomoléculaire complexe

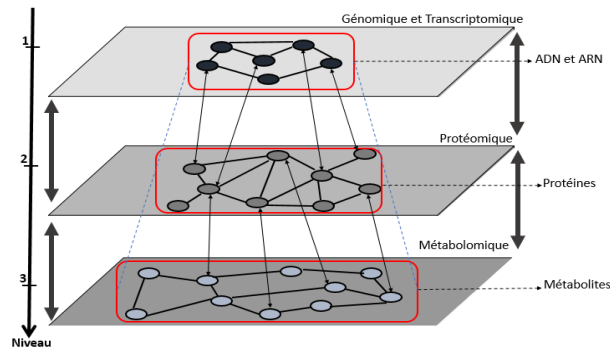


FIG. 1 – Représentation multi-échelle du réseau biomoléculaire.

La cellule est composée d'éléments moléculaires qui interagissent les uns avec les autres physiquement, fonctionnellement et logiquement formant ainsi un réseau biomoléculaire. Selon le type de ces éléments et de leurs interactions, nous distinguons trois types de réseaux qui ont fait l'objet de la plupart des études récentes [Lombardi et Hörnquist \(2007\)](#), [Karp \(2010\)](#) et [Mones et al. \(2012\)](#) (Figure 1) :

1. Le réseau génétique qui modélise les interactions entre à environ 21 000 gènes (ADNs et ARNs). Il est représenté par un graphe dirigé où les nœuds représentent les gènes et les arcs modélisent le type de régulation (activation ou inhibition) d'un gène sur un autre.
2. Le réseau protéine-protéine qui modélise les interactions ayant lieu entre environ 80 000 protéines. Ce réseau est représenté par un graphe non orienté où les nœuds sont les protéines et les arêtes non dirigées modélisent les liaisons entre ces nœuds. Ces types d'interactions dépendent de la fonction des protéines.
3. Le réseau métabolique modélise les réactions du métabolisme c'est à dire les réactions biochimiques entre les environ 42 000 métabolites réactants et produits. Un réseau métabolique est représenté par un graphe dont les nœuds sont des métabolites (réactants ou produits) et les arcs représentent le type de réaction biochimique transformant les réactants en des produits, elles sont marquées par le coefficient stœchiométrique du métabolite dans la réaction.

Ainsi, un réseau biomoléculaire peut être représenté par un graphe noté généralement par le quadruplet suivant :

$$G = (M, I, c, v)$$

Avec :

- $M$  : l'ensemble des molécules formant le réseau représente les nœuds du graphe  $G$  : défini par un ensemble fini de sommets  $M = \{m_1, m_2, \dots, m_n\}$ . Nous distinguons une partition tripartite de nœuds : les ensembles  $M_G$  des gènes (ADNs et ARNs),  $M_P$

des protéines et  $M_M$  des métabolites. Nous pourrions ainsi écrire :

$$M = M_G \cup M_P \cup M_M$$

- $I$  : l'ensemble des interactions entre les molécules formant le réseau. Représente les arcs du graphe  $G$  : défini par un ensemble fini d'arcs ou d'arêtes  $I = \{I_1, I_2, \dots, I_m\}$ . L'arc  $I = (m_i, m_j)$ , (avec  $m_i, m_j \in M$ ) qui part de  $m_i$  (origine) et arrive à  $m_j$  (destination) sera également noté  $m_i \rightarrow m_j$ . La partition des nœuds du graphe induit aussi une partition des arcs en divers types d'interactions :
  - . trois interactions entre des composants moléculaires de même type : les interactions entre les gènes noté  $I_{GG}$  qui modélise le type de régulation entre les gènes (activation ou inhibition),  $I_{PP}$  qui modélise les associations stables ou transitoires entre les protéines et  $I_{MM}$  modélisant les interactions ayant lieu entre les métabolites (type de réaction chimique entre réactants et produits).
  - . quatre interactions (parmi les 6 possibles) entre noeuds de réseaux différents :  $I_{GP}$  qui modélise les régulation des gènes sur les protéines,  $I_{PG}$  qui modélise l'influence des protéines sur les gènes à travers le facteur de transcription,  $I_{PM}$  représente les enzymes intervenant dans les réactions chimiques des métabolites (catalyse ou hydrolyse),  $I_{MP}$  modélise l'influence des métabolites sur les protéines.
  - . deux interactions  $I_{GM}$  et  $I_{MG}$  ne sont pas prises en compte car il n'y a pas d'interaction directe entre les gènes et les métabolites.

$$I = I_{GG} \cup I_{PP} \cup I_{MM} \cup I_{GP} \cup I_{PM} \cup I_{MP} \cup I_{PG}$$

- $c$  : la fonction qui associe à chaque nœud (molécule) du graphe la valeur de sa concentration et les deux seuils (minimal et maximal) qui la déclenche. Ces seuils que l'on note  $S_{(M_i)_{min}}$  et  $S_{(M_i)_{max}}$  représentent respectivement les limites de concentration minimale et maximale entre lesquels l'interaction suivante va se déclencher.
- $v$  : la fonction qui associe à chaque arc une étiquette représentant la nature de l'interaction (hydrolyse, régulation positive, régulation négative, etc.).

L'état d'un réseau biomoléculaire est défini comme l'ensemble des valeurs représentant les niveaux de concentration dans ses noeuds à un instant  $t$ . L'état du réseau  $G$  à l'instant  $t$  est modélisé par le vecteur d'état suivant :

$$C_{(G)_t} = (c_1, c_2, \dots, c_n)_t$$

où  $c_i$  (avec  $\{i \in [1..n]\}$ ) est la valeur de concentration dans le nœud  $m_i \in M$ .

Une transition d'état dans le réseau peut se produire soit par un stimulus interne (augmentation de la concentration d'une molécule, par exemple) soit par un stimulus externe (la prise de médicament par exemple).

### 3 Suite de nos travaux

L'objectif général de cette thèse est de trouver un ensemble de stimulus externes optimal à appliquer pendant un intervalle de temps pré-fixé pour faire évoluer le réseau de son état

courant à un autre état souhaité. Notre approche innovante est basée sur l'utilisation conjointe des technologies sémantiques, d'optimisation combinatoire et de simulation.

Avec ce but, nos travaux futurs se poursuivront par le développement d'une plate-forme pour étudier les transitions des réseaux biomoléculaires d'un état quelconque à un état spécifique, basée sur trois modules :

1. Le module ontologique : Ce module utilise des technologies sémantiques pour générer de nouvelles connaissances inférées (la découverte de nouvelles associations sémantiques entre les molécules) afin d'affiner l'étude des transitions du réseau. Ce modèle prend en entrée l'ensemble des informations natives (état du réseau ses transitions sous forme de valeurs et de paramètres) introduites par l'expert et fournit au final le réseau inféré constitué de transition d'états natifs et inférés. Cet enrichissement par des méta-données et de nouvelles connaissances facilitera les prises de décision grâce à une gestion performante de la connaissance [Zanni-Merk \(2014\)](#).
2. Le module de simulation : Ce module reproduira au cours du temps le comportement dynamique de chaque composant du réseau. Ce simulateur adoptera le formalisme DEVS Discrete Event Specification de [Zeigler et al. \(2000\)](#).
3. Le module d'optimisation : Avec ce module, l'application des algorithmes d'optimisation combinatoire pour fournir un ensemble de séquences de transitions offrant le meilleur pilotage du réseau d'un état à un autre, tout en décrivant les changements de valeurs ayant lieu à chaque composant du réseau.

## Références

- Karp, G. (2010). *Biologie cellulaire et moléculaire : Concepts and experiments*. De Boeck Supérieur.
- Lombardi, A. et M. Hörnquist (2007). Controllability analysis of networks. *Phys. Rev. E* 75, 056110.
- Mones, E., L. Vicsek, et T. Vicsek (2012). Hierarchy measure for complex networks. *PLoS ONE* 7(3), e33799.
- Wu, F.-X., L. Wu, J. Wang, J. Liu, et L. Chen (2014). Transittability of complex networks and its applications to regulatory biomolecular networks. *Scientific reports* 4.
- Zanni-Merk, C. (2014). *Knowledge technologies for problem solving in Engineering*. Mémoire d'Habilitation À Diriger des Recherches. Université de Strasbourg.
- Zeigler, B. P., H. Praehofer, et T. G. Kim (2000). *Theory of modeling and simulation : integrating discrete event and continuous complex dynamic systems*. Academic press.

## Summary

The "transittability" of a complex molecular network expresses the idea of steering this network from an undesired state (usually associated with illness) to a desired state. In this article, we present a formalization of complex biomolecular networks and we propose a platform based on semantic technologies for optimizing the transittability of these biomolecular networks.