

THÈSE DE DOCTORAT DE

NANTES UNIVERSITÉ

ÉCOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *INFO*

Par

Sophie LE BARS

Hybrid, logical and linear modelling to predict *in silico* the effect of disturbances on metabolism

Thèse présentée et soutenue à «Nantes», le «9 décembre 2022»

Unité de recherche : LS2N - Laboratoire des sciences et du numérique de Nantes

Rapporteurs avant soutenance :

Alexander BOCKMAYR Professeur, Freie Universität Berlin, Berlin, Allemagne
Hervé ISAMBERT Directeur de Recherche CNRS, Institut Curie, Paris, France

Composition du Jury :

Attention, en cas d'absence d'un des membres du Jury le jour de la soutenance, la composition du jury doit être revue pour s'assurer qu'elle est conforme et devra être répercutée sur la couverture de thèse

Président : Prénom NOM, Fonction et établissement d'exercice (*à préciser après la soutenance*)
Examineurs : Alexander BOCKMAYR, Professeur, Freie Universität Berlin, Berlin, Allemagne
Hervé ISAMBERT, Directeur de Recherche CNRS, Institut Curie, Paris, France
Anne SIEGEL, Directrice de Recherche CNRS, Irisa, Rennes, France
Maxime FOLSCHETTE, Maître de conférences, Centrale Lille, Lille, France
Dir. de thèse : Jérémie BOURDON, Professeur, Nantes Université, Nantes, France
Co-enc. de thèse : Carito GUZIOLOWSKI, Maitresse de conférences, Centrale Nantes, Nantes, France

Titre : Modélisation hybride, logique et linéaire pour prédire *in silico* l'effet des perturbations sur le métabolisme

Mot clés : Modélisation, Answer Set Programming, Intégration de réseaux

Résumé : Les perturbations induites par une maladie, un traitement ou encore un stress environnemental affectent les organismes vivants de diverses manières. Ainsi, l'expression de certains gènes sera impactée, ce qui se répercutera sur ses produits (protéines, ARNm). Ces perturbations se propagent également via les interactions que peuvent avoir les gènes les uns avec les autres. L'ensemble de ces interactions forme le réseau de régulation, un objet important dans ces études. D'autre part, certaines protéines, appelées enzymes, ont un rôle de catalyseur des réactions biochimiques qui ont lieu au sein des organismes. L'ensemble des réactions biochimiques forme le réseau métabolique, un second objet important. Ainsi, une perturbation va impacter le réseau de régulation mais aussi le réseau métabolique puisqu'ils sont interconnectés, via les enzymes notamment. L'objectif principal de ma thèse est d'étudier l'impact d'une perturbation sur un or-

ganisme en intégrant le réseau de régulation au réseau métabolique. J'ai apporté deux contributions dans ce sens. La première est une comparaison d'une approche logique à une approche bayésienne pour savoir quelle stratégie de modélisation est la plus adaptée pour étudier les impacts des perturbations sur de grands réseaux de régulations. J'en ai déduit que bien qu'elle soit une bonne candidate, l'approche logique présente des limites de par ses prédictions qualitatives en matière d'intégration. La seconde contribution découle de ces limites, j'ai développé une méthode originale basée sur l'Answer Set Programming, MajS, proposant une prédiction plus fine de l'effet d'une perturbation sur le réseau de régulation. Ce travail ouvre la porte à de nombreuses perspectives comme une meilleure intégration des effets des perturbations au niveau du réseau métabolique et une application à d'autres organismes d'étude.

Title: Hybrid, logical and linear modeling to predict *in silico* the effect of disturbances on metabolism

Keywords: Computer modelling, Answer Set Programming, Networks integration

Abstract: Disturbances induced by disease, treatment or environmental stress affect living organisms in various ways. Thus, the expression of certain genes will be impacted, affecting its products (proteins, mRNA). These disturbances are also propagated via the interactions that genes can have. These interactions form the regulatory network, an essential object in these studies. On the other hand, specific proteins, called enzymes, act as catalysts for the biochemical reactions occurring within organisms. All the biochemical reactions form the metabolic network, a second important object. Thus, a disturbance will impact the regulatory and metabolic networks since they are interconnected, via enzymes in particular. My thesis's main objective is to study a disturbance's impact on an organism by integrating the regulatory network into the metabolic

network. I have made two contributions in this direction. The first compares a logical approach to a Bayesian approach to determine which modelling strategy is the most suitable for studying the impacts of disturbances on large regulatory networks. Although it is a good candidate, I deduced that the logical approach has limitations with its qualitative predictions regarding integration. The second contribution stems from these limits; I have developed an original method based on Answer Set Programming, MajS, offering a more refined prediction of the effect of a disturbance on the regulation network. This work opens the door to many perspectives, such as better integration of the effects of disturbances at the metabolic network level and an application to other organisms of study.