# THÈSE DE DOCTORAT DE

L'ÉCOLE CENTRALE DE NANTES
**C**OMUE **U**NIVERSITE **B**RETAGNE **L**OIRE

Ecole Doctorale N°601
*Mathèmatique et Sciences et Technologies de l'Information et de la Communication*
Spécialité : *Informatique*

Par

« **Xinwei CHAI** »

« **Reachability Analysis and Revision of Dynamics of Biological Regulatory Networks** »

«Analyse d'accessibilité et révision de la dynamique dans les réseaux de régulations biologiques»

**Thèse présentée et soutenue à** L'ÉCOLE CENTRALE DE NANTES, le 24 mai, 2019

**Rapporteurs avant soutenance :**

| | | |
|---|---|---|
| Gilles Bernot | Professeur des universités | Université Côte d'Azur, Sophia Antipolis |
| Pascale Le Gall | Professeur des universités | CentraleSupélec, Gif sur Yvette |

**Composition du jury :**

| | | | |
|---|---|---|---|
| Président : | Béatrice Duval | Professeur des universités | Université d'Angers |
| Examinateurs : | Gilles Bernot | Professeur des universités | Université Côte d'Azur, Sophia Antipolis |
| | Pascale Le Gall | Professeur des universités | CentraleSupélec, Gif sur Yvette |
| | Morgan Magnin | Professeur des universités | École Centrale de Nantes |
| | Loïc Paulevé | Chargé de recherche | Université de Bordeaux |
| | Olivier Roux | Professeur des universités | École Centrale de Nantes |
| Dir de thèse : | Olivier Roux | Professeur des universités | École Centrale de Nantes |
| Co-dir. de thèse : | Morgan Magnin | Professeur des universités | École Centrale de Nantes |

# Titre: Analyse d'accessibilité et révision de la dynamique dans les réseaux de régulations biologiques

**Mots clés :** bioinformatique, model checking, heuristique, révision de modèles

**Resumé :** Les systèmes concurrents sont un bon choix pour ajuster les données et analyser les mécanismes sous-jacents pour leur sémantique simple mais expressive. Cependant, *l'apprentissage et l'analyse* de tels systèmes concurrents sont difficiles pour ce qui concerne les calculs. Lorsqu'il s'agit de grands ensembles de données, les techniques les plus récentes semblent insuffisantes, que ce soit en termes d'efficacité ou de précision. Ici, nous proposons un cadre de modélisation raffiné ABAN (Asynchronous Binary Automata Network) et développons des outils pour analyser l'atteignabilité : PermReach (Reachability *via* Permutation search) et ASPReach (Reachability *via* Answer Set Programming). Nous proposons ensuite deux méthodes de construction et d'apprentissage des modèles: CRAC (Completion *via* Reachability And Correlations) et M2RIT (Model Revision *via* Reachability and Interpretation Transitions) en utilisant des données continues et discrètes pour s'ajuster au modèle et des propriétés d'accessibilité afin de contraindre les modèles en sortie.

# Title: Reachability Analysis and Revision of Dynamics of Biological Regulatory Networks

**Keywords :** bioinformatics, model checking, heuristics, model revision

**Abstract :** Concurrent systems become a good choice to fit the data and analyze the underlying mechanics for their simple but expressive semantics. However, *learning* and *analyzing* such concurrent systems are computationally difficult. When dealing with big data sets, the state-of-the-art techniques appear to be insufficient, either in term of efficiency or in term of precision. In this thesis, we propose a refined modeling framework ABAN (Asynchronous Binary Automata Network) and develop reachability analysis techniques based on ABAN: PermReach (Reachability *via* Permutation search) and ASPReach (Reachability *via* Answer Set Programming). Then we propose two model learning/constructing methods: CRAC (Completion *via* Reachability And Correlations) and M2RIT (Model Revision *via* Reachability and Interpretation Transitions) using continuous and discrete data to fit the model and using reachability properties to constrain the output models.

# ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my advisors Prof. Olivier ROUX and Prof. Morgan MAGNIN for the continuous support of my Ph.D. study for their patience, ideas and contribution of time. This kind support also came to my personal life which helped me to regain motivation and carry on the research when I encountered at the same time academy and mental difficulties. I especially appreciate their tolerance allowing me to explore on my will even the outcome was not satisfying.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Gilles BERNOT, Prof. Pascal LE GALL, Prof Béatrice DUVAL and Dr. Loïc PAULEVÉ, for their patient reading, insightful comments and encouragement, but also for the questions which incented me to widen my research from various perspectives.

My sincere thanks also goes to my kind team members. Dr. Emna BEN ABDALLAH, Dr. Maxime FOLSCHETTE, Samuel BUCHET and Dr. Loïc PAULEVÉ provided me an opportunity to work with them even after their graduation. Without their precious but abundant support, it would not be possible for me to conduct this research. I would like to especially acknowledge Tony RIBEIRO Sensei for his kindness, enthusiasm, intensity and genius ideas which redirected and accelerated my Ph.D. study.

I gratefully acknowledge the funding provided by China Scholarship Council (CSC) that made my Ph.D. work possible.

My life in Nantes is enriched by my warm-hearted neighbors Geneviève ROCHE and the family of MARCHAND. Also I thank my friends in École Centrale de Nantes for all the fun we have had in the last four years.

Last but not the least, I would like to thank my family for all their love and encouragement. For my parents who raised me with curiosity in science and supported me in all my pursuits. For my uncle who talked with me intimately in my sadness and confusion. For my girlfriend who accompanied me spiritually everyday across the ocean.

Thank you.

Xinwei Chai
École Centrale de Nantes
May 2019