

# Modélisation dynamique du développement embryonnaire humain

## 1 Contexte

Le projet s'inscrit dans le contexte de l'étude et de la modélisation du développement embryonnaire humain. Grâce aux récentes avancées technologiques en analyse de données biologiques, il est désormais possible de mesurer l'expression des gènes au niveau des cellules individuelles (scRNAseq) d'un organisme. En analysant ces données avec des approches classiques, nous pouvons identifier des groupes de gènes qui se comportent de manière similaire à différents stades du développement embryonnaire. Cependant, nous cherchons à aller plus loin en modélisant informatiquement ce processus.

Actuellement, nous avons implémenté une méthode [1] utilisant des réseaux booléens pour modéliser les interactions entre les gènes à deux stades de développement clés. Ces réseaux booléens nous permettent de distinguer ces stades, d'identifier les gènes impliqués, et ainsi de mieux comprendre les mécanismes sous-jacents à la transition entre ces stades.

## 2 Problématique

Ce projet vise à introduire la dimension de la dynamique dans notre modélisation. Il se décline en deux objectifs principaux :

**Objectif 1** Étendre la modélisation à plus de deux stades de développement et inférer des modèles représentant la dynamique du développement embryonnaire pour comprendre la séquence temporelle des stades. Dans un premier temps, le travail consistera à reprendre et adapter un programme logique préliminaire qui utilise des notions de *l'analyse formelle de concepts* [2] pour extraire de comportements d'expression (booléens) différents et relatifs à un ensemble de gènes spécifique. Cette approche pourra être enrichie avec des outils, comme **Palantir** [3], qui ordonnent les cellules avec des *manifolds* (ou variétés, technique d'analyse du *big data*) pour obtenir l'évolution de l'expression des gènes au cours du développement embryonnaire. Enfin, **caspo-ts** [4] sera utilisé pour inférer des réseaux booléens représentant cette dynamique.

**Objectif 2** Si les résultats de l'objectif 1 sont satisfaisants et que le temps le permet, explorer d'autres outils de modélisation de la dynamique cellulaire tels que **BoNesis** [5], adaptés à nos données, pour une approche alternative de la modélisation du développement embryonnaire. Une comparaison entre les différentes approches pourra être réalisée, et des consultations avec des biologistes experts pourront être organisées pour une analyse approfondie des résultats obtenus.

## 3 Paradigme

Le paradigme de programmation central pour ce projet est l'*Answer Set Programming* (ASP) [6], également connu sous le nom de "programmation par ensemble réponses". Modéliser en ASP se fait à travers des programmes logiques qui utilisent des prédicats du premier ordre. Ce système est particulièrement adapté à la modélisation de problèmes de recherche combinatoire impliquant des variables discrètes, tels que des problèmes liés à des puzzles, des graphes, la planification, etc.

**Équipe d'accueil** LS2N Laboratoire des Sciences du Numérique de Nantes. Équipe COMBI - Combinatoire et Bioinformatique. Université de Nantes. **Encadrant·e·s** Carito Guziolowski (carito.guziolowski@ec-nantes.fr), Mathieu Bolteau (mathieu.bolteau@univ-nantes.fr). Pour plus de précisions sur le sujet, merci de nous contacter par email.

**Environnements et technologies utilisées** Clasp<sup>1</sup> [7], Python, tout autre package et librairie jugés utiles.

---

1. <https://potassco.org/clasp/>

**Mots clés** Intelligence artificielle, graphes, combinatoire, programmation logique, développement embryonnaire humain

## Références

- [1] M. Bolteau, J. Bourdon, L. David, and C. Guziolowski, “Inferring Boolean Networks from Single-Cell Human Embryo Datasets,” Aug. 2023, working paper or preprint. [Online]. Available : <https://hal.science/hal-04206397>
- [2] B. Ganter and R. Wille, *Formal Concept Analysis : Mathematical Foundations*. Berlin/Heidelberg : Springer, 1999.
- [3] M. Setty, V. Kiseliouas, J. Levine, A. Gayoso, L. Mazutis, and D. Pe’er, “Characterization of cell fate probabilities in single-cell data with Palantir,” *Nature Biotechnology*, vol. 37, no. 4, pp. 451–460, Apr. 2019. [Online]. Available : <https://www.nature.com/articles/s41587-019-0068-4>
- [4] M. Razzaq, L. Paulevé, A. Siegel, J. Saez-Rodriguez, J. Bourdon, and C. Guziolowski, “Computational discovery of dynamic cell line specific Boolean networks from multiplex time-course data,” *PLoS Comput Biol*, vol. 14, no. 10, p. e1006538, 10 2018.
- [5] S. Chevalier, V. Noël, L. Calzone, A. Zinovyev, and L. Paulevé, “Synthesis and simulation of ensembles of boolean networks for cell fate decision,” in *Computational Methods in Systems Biology*, A. Abate, T. Petrov, and V. Wolf, Eds. Cham : Springer International Publishing, 2020, pp. 193–209.
- [6] C. Baral, *Knowledge Representation, Reasoning and Declarative Problem Solving*. Cambridge : Cambridge University Press, 2003. [Online]. Available : <https://doi.org/10.1017/CBO9780511543357>
- [7] M. Gebser, B. Kaufmann, and T. Schaub, “Conflict-driven answer set solving : From theory to practice,” *Artif. Intell.*, vol. 187, pp. 52–89, 2012.