

institut**Curie**

A logical approach to identify Boolean networks modeling cell differentiation

Stéphanie Chevalier⁽¹⁾, Mihaly Koltai⁽²⁾, Andrei Zinovyev⁽²⁾, Christine Froidevaux⁽¹⁾, Loïc Paulevé⁽¹⁾

(1) Laboratoire de Recherche en Informatique (LRI), Université Paris Sud - Paris Saclay, CNRS, France (2) U900 Cancer et génome : bioinformatique, biostatistiques et épidémiologie, Institut Curie, Inserm, MINES ParisTech, France

issue : What are the genes interactions that lead to differentiation?	approach : Inference of Boolean networks compatible with differentiation data
A priori knowledge	Qualitative method
<i>Fluence graph</i> from databases and expert knowledge ·	Allow a weak influence of the biological data approximation

Influence graph from databases and expert knowledge : **Prior Knowledge Network (PKN)**



- no information on genes cooperation :

R IIIIIUEIILE UI LIE DIVIUGILAI UALA to focus on the essential dynamics features.

Boolean network (BN) with n genes :

 $F = (f_1, ..., f_n), \forall i \in \{1, ..., n\}, f_i : \mathbb{B}^n \to \mathbb{B} \ (\mathbb{B} = \{0, 1\}),$

with f_i the target value of the ith gene

Boolean network dynamic : transition graph



Nodes : states

(active/inactive status of all the genes)

Edges : asynchronous updates (one gene is updated)

Differentiation data

Experimental measurements : time series of differentiation process

Measurements for only a few genes,

at first in undifferentiated and then in differentiated cells.

Example :

- T0 : undifferentiated cell
- gene1 inactive
- gene2 active

T1>T0 : differentiated cell type 1 gene1 active • gene2 active

T1>T0 : differentiated cell type 2

Boolean networks inference

Objective : exhaustive list of compatible BN

- to study the inferred BN variability (graph influence topology & nodes importance)
- to quantify the data informativeness (data relevance to infer BN)

Direct approach :



Transition graph

* **Attractor :** minimal set of states without outgoing transition (long-term dynamics)

gene1 active • gene2 inactive

Data interpretation (hypotheses) :

- **Reachability** : state T_{i+1} reachable from T_i
- **Specialization** : differently differentiated states have no common descendant
- **Stability** : stable differentiated states are in attractors^{*}
- **Representativeness** : proportion of states compatible with the observations

How to infer Boolean networks that respect these properties?

Conclusion & **Perspectives**





Constraints impact :

Exhaustive model inference for **differentiation data**

- thanks to **constraints** fitting the experimental conditions (fixed points already implemented)
- allows data informativeness quantification

Coming work :

- Implementation of negative reachability conditions
- Generalization of the method to mutation data (perturbed differentiation) data) and **quantitative** mutation data (fates proportion w.r.t. mutation)
- **Predict mutations combinations** that trigger fates

- necessary condition on <u>reachability</u> (Caspots^[1])
- necessary and sufficient condition on a stability hypothesis :





References

[1] Max Ostrowski, Loïc Paulevé, Torsten Schaub, Anne Siegel, Carito Guziolowski. Boolean Network Identification from Perturbation Time Series Data combining Dynamics Abstraction and Logic Programming. BioSystems, Elsevier, 2016

[2] Qiu X, Ding S, Shi T. From Understanding the Development Landscape of the Canonical Fate-Switch Pair to Constructing a Dynamic Landscape for Two-Step Neural Differentiation. PLOS ONE 7(12):e49271, 2012

