Simulation and synthesis of formal models of biological systems

Athénaïs Vaginay
Post-doc in MTV (dep. M2F), with Loïc Paulevé

Sémidoc @ LaBRI, Bdx — 2024 July 19

Outline

1. My curriculum :)))
2. A light intro to systems biology and formal models in general
3. Formal methods for the simulation of reaction networks
4. Formal methods for the simulation of Boolean networks
5. Formal methods for the synthesis of Boolean networks
6. Conclusion

Curriculum

Me and my background

Currently (for two more weeks): post-doc in MTV, under the supervision of Loïc Paulevé.

- Medical studies, 1 year
- Biology, licence
- Bio-info, master + "ingénieure d'études" a few months
- Theoretical systems biology ($\sim$ computer science), PhD, postdoc, ...

Reach out: room B353, `athenais.vaginay@labri.fr`

My first time in this amphi: in 2021 (CMSB conference).

Introduction

Systems biology

Formal modelling and reasoning about biological systems

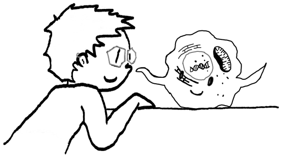
A model = an abstract representation (abbreviated and convenient) of the reality (more complex and detailed)

A set species of species of interest genes, proteins, cells, animals. . .

Questions

How does the system evolve?

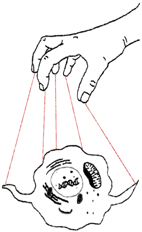
Is the population of some cell type stable over time?



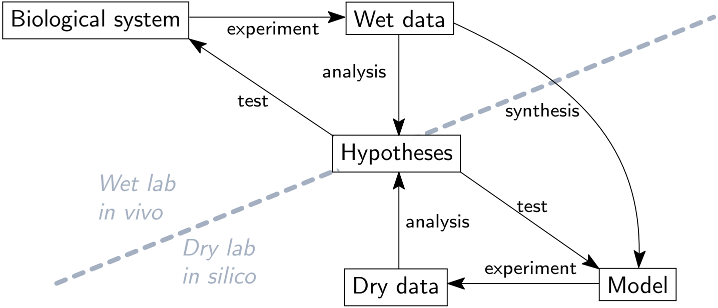
How to control the system?

Cure a pathological system

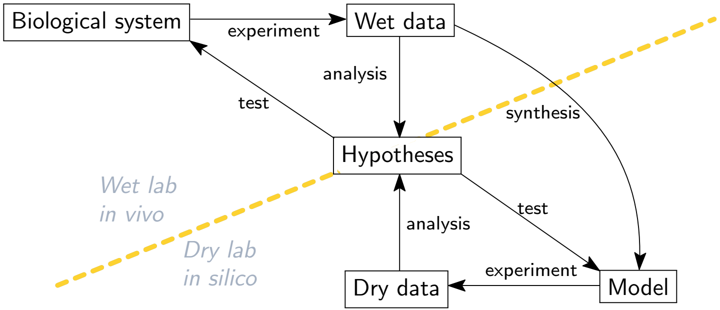
Produce more of some species of interest



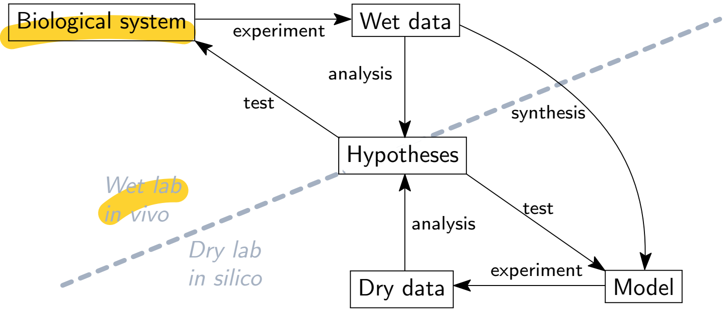
The workflow of system biology [Kohl et al., 2010]



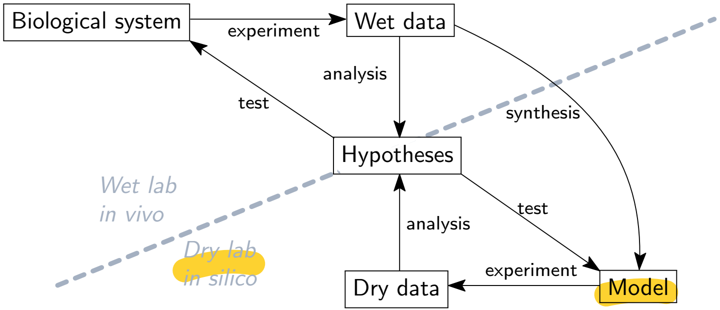
The workflow of system biology [Kohl et al., 2010]



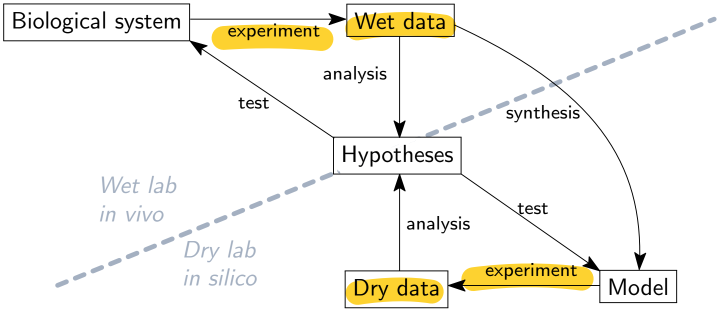
The workflow of system biology [Kohl et al., 2010]



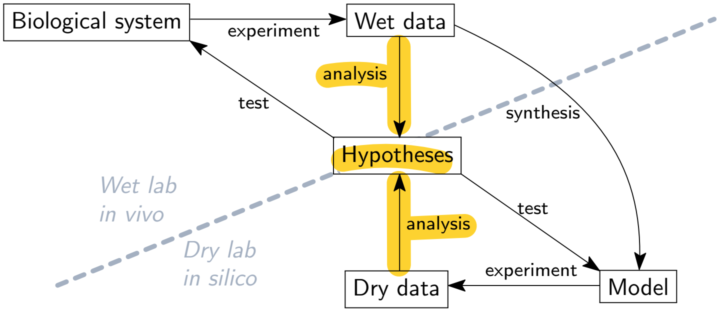
The workflow of system biology [Kohl et al., 2010]



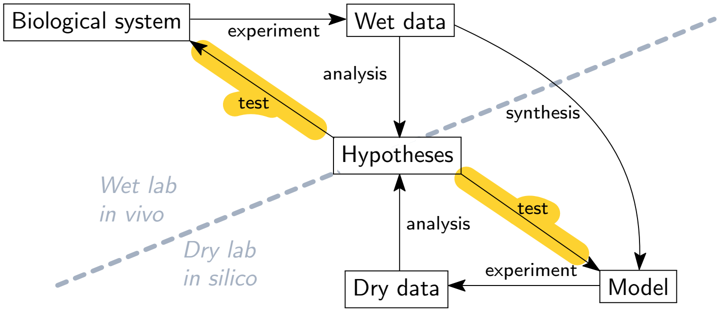
The workflow of system biology [Kohl et al., 2010]



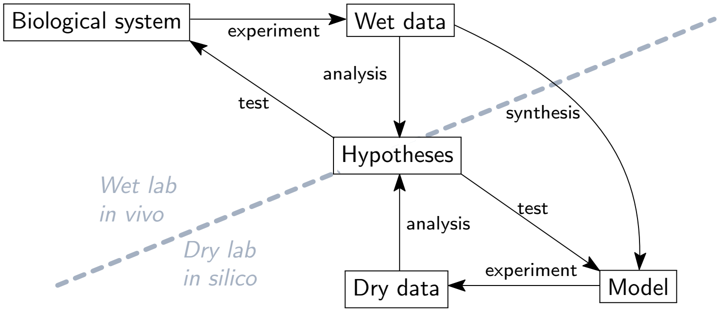
The workflow of system biology [Kohl et al., 2010]



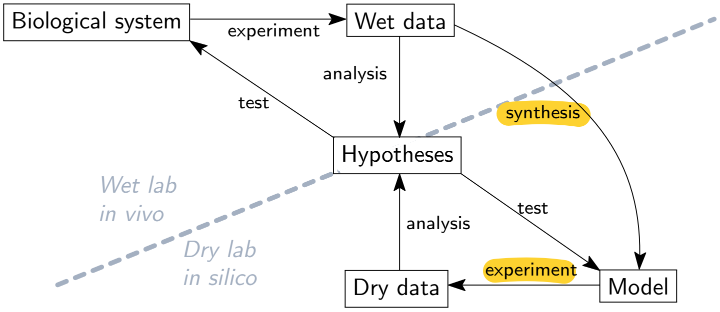
The workflow of system biology [Kohl et al., 2010]



The workflow of system biology [Kohl et al., 2010]



The workflow of system biology [Kohl et al., 2010]



Synthesis:

- from available knowledge and data about the structure and the dynamics
- parameter fitting task find models that optimise some criteria

Experiment: e.g. simulation = execution of the model

A zoo of modelling approaches

Reaction network

continuous time Markov chain

ODEs

statistical models

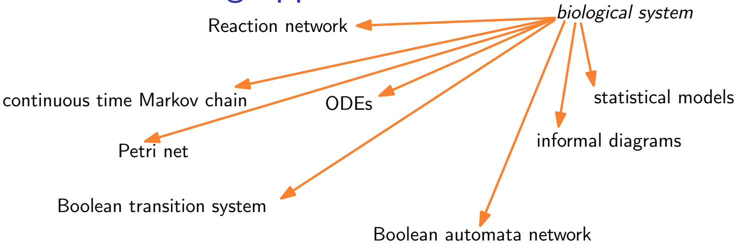
informal diagrams

Petri net

Boolean transition system

Boolean automata network

A zoo of modelling approaches



informal **abstractions**

A zoo of modelling approaches

Reaction network

continuous time Markov chain

ODEs

statistical models

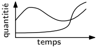
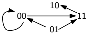
informal diagrams

Petri net

Boolean transition system

Boolean automata network

Model = rules A → B

Reaction networks	**Boolean networks**
Transformation of reactants into products	Influence of regulators on regulees
Differential equations	Boolean functions (propositional logics)
$\dot{B} = f(A)$	$B = f(A)$
	
Time series	Transition graph
Continuous time and values	Discrete time, boolean values

Formal methods for the simulation of reaction networks

Static analysis of a reaction network

$$\mathcal{R} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i \mid i = 1 \dots m\}$$

reaction, reactants, products, kinetics

Example

$$\mathcal{S} = \{\mathsf{A}, \mathsf{B}, \mathsf{C}\}$$

$$\mathcal{R}_1 : \mathsf{A} + \mathsf{B} \xrightarrow{e_1} 2 \times \mathsf{C}$$

$$\mathcal{R}_2 : \mathsf{C} \xrightarrow{e_2} \mathsf{B}$$

Static analysis = derive correct conclusions about the dynamics without having to actually simulate the model.

Static analysis of a reaction network

$\mathcal{R} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i \mid i = 1 \ldots m\}$

reaction, reactants, products, kinetics

Example

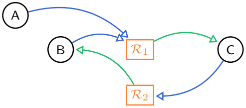
$\mathcal{S} = \{\mathsf{A}, \mathsf{B}, \mathsf{C}\}$

$\mathcal{R}_1 : \mathsf{A} + \mathsf{B} \xrightarrow{e_1} 2 \times \mathsf{C}$

$\mathcal{R}_2 : \mathsf{C} \xrightarrow{e_2} \mathsf{B}$

Reaction graph

$(\mathcal{S} \cup \mathcal{R}, E \subseteq (\mathcal{S} \times \mathcal{R}) \cup (\mathcal{R} \times \mathcal{S}))$



Static analysis = derive correct conclusions about the dynamics without having to actually simulate the model.

Static analysis of a reaction network

$\mathcal{R} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i \mid i = 1 \ldots m\}$
reaction, reactants, products, kinetics

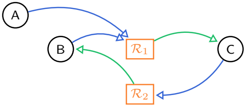
Reaction graph
$(\mathcal{S} \cup \mathcal{R}, E \subseteq (\mathcal{S} \times \mathcal{R}) \cup (\mathcal{R} \times \mathcal{S}))$

Example

$\mathcal{S} = \{\mathsf{A}, \mathsf{B}, \mathsf{C}\}$

$\mathcal{R}_1 : \mathsf{A} + \mathsf{B} \xrightarrow{e_1} 2 \times \mathsf{C}$

$\mathcal{R}_2 : \mathsf{C} \xrightarrow{e_2} \mathsf{B}$



$\rightsquigarrow$ A can only decrease

Static analysis = derive correct conclusions about the dynamics without having to actually simulate the model.

Abstract simulation of a reaction network

Reaction network

$$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\ldots m}$$

Abstract simulation of a reaction network

Reaction network

$$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\ldots m}$$

$$A + B \xrightarrow{e} 2C$$

Abstract simulation of a reaction network

Reaction network

$$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\ldots m}$$

$$A + B \xrightarrow{e} 2C$$

Continuous Time Markov chain

Ordinary Differential Equations

Petri net

Boolean transition system

Abstract simulation of a reaction network

Reaction network

$$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\ldots m}$$

$$A + B \xrightarrow{e} 2C$$

Continuous Time Markov chain

Ordinary Differential Equations

continuous time, continuous values

$$\dot{A} = \dot{B} = -e; \dot{C} = 2e$$

Petri net

Boolean transition system

Abstract simulation of a reaction network

Reaction network

$$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\ldots m}$$

$$A + B \xrightarrow{e} 2C$$

Continuous Time Markov chain

continuous time, discrete values

$p(e)$: A- -; B- -; C+=2

Ordinary Differential Equations

continuous time, continuous values

$\dot{A} = \dot{B} = -e; \dot{C} = 2e$

Petri net

Boolean transition system

Abstract simulation of a reaction network

Reaction network

$$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\ldots m}$$

$A + B \xrightarrow{e} 2C$

Continuous Time Markov chain

continuous time, discrete values

$p(e)$: A- -; B- -; C+=2

Ordinary Differential Equations

continuous time, continuous values

$\dot{A} = \dot{B} = -e; \dot{C} = 2e$

Petri net

discrete time, discrete values

A- -; B- -; C+=2

Boolean transition system

Abstract simulation of a reaction network

Reaction network

$$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\ldots m}$$

$$A + B \xrightarrow{e} 2C$$

Continuous Time Markov chain

continuous time, discrete values

$p(e)$: A- -; B- -; C+=2

Ordinary Differential Equations

continuous time, continuous values

$\dot{A} = \dot{B} = -e$; $\dot{C} = 2e$

Petri net

discrete time, discrete values

A- -; B- -; C+=2

Boolean transition system

discrete time, boolean values

A, B `in` $\{0,1\}$; C `in` $\{1\}$

$$\begin{tikzcd}[row sep=tiny]
 & 101 \\
110 \arrow[ur] \arrow[r] \arrow[dr] \arrow[ddr] & 011 \\
 & 111 \\
 & 001
\end{tikzcd}$$

Abstract simulation of a reaction network

Reaction network

$\boldsymbol{\mathcal{R}} = \{\mathcal{R}_i : R_i \xrightarrow{e_i} P_i\}_{i=1\dots m}$

$A + B \xrightarrow{e} 2C$

Continuous Time Markov chain

continuous time, discrete values

$p(e)$: A- -; B- -; C+=2

Ordinary Differential Equations

continuous time, continuous values

$\dot{A} = \dot{B} = -e;\ \dot{C} = 2e$

Petri net

discrete time, discrete values

A- -; B- -; C+=2

Boolean transition system

discrete time, boolean values

A, B `in` $\{0,1\}$; C `in` $\{1\}$

$110 \to 101$, $110 \to 011$, $110 \to 111$, $110 \to 001$

Formal abstraction between two **semantics**

Abstract simulation = derive correct conclusions using a simpler simu. of the model [Cousot, Cousot, 1977], [Fages, Soliman, 2008a]

Formal methods for the simulation of Boolean networks

The dynamics of a Boolean network using a SAT solveur

A BN f is a function $\mathbb{B}^n \to \mathbb{B}^n$ usually expressed in propositional logics.

The dynamics of a Boolean network using a SAT solveur

A BN f is a function $\mathbb{B}^n \to \mathbb{B}^n$ usually expressed in propositional logics.

Example

$$x_1' = (x_1 \wedge x_2 \wedge x_3) \wedge x_2' = (x_1 \vee x_3) \wedge x_3' = ((x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3))$$

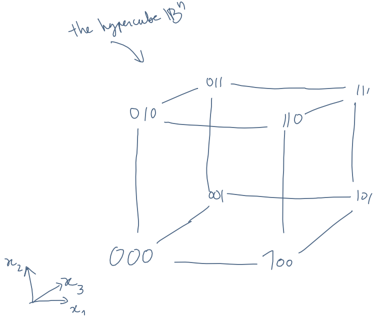
The dynamics of a Boolean network using a SAT solveur

A BN f is a function $\mathbb{B}^n \to \mathbb{B}^n$ usually expressed in propositional logics.

Transition graph: $G = (\mathbb{B}^n, E \subseteq \mathbb{B}^n \times \mathbb{B}^n)$. Is $(x, x') \in G$?

Example

$$x_1' = (x_1 \wedge x_2 \wedge x_3) \wedge x_2' = (x_1 \vee x_3) \wedge x_3' = ((x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3))$$

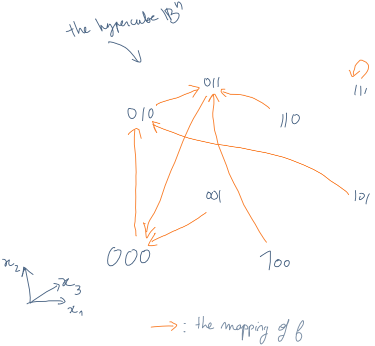


The dynamics of a Boolean network using a SAT solveur

A BN f is a function $\mathbb{B}^n \to \mathbb{B}^n$ usually expressed in propositional logics.
Transition graph: $G = (\mathbb{B}^n, E \subseteq \mathbb{B}^n \times \mathbb{B}^n)$. Is $(x, x') \in G$?

Example

$$x_1' = (x_1 \wedge x_2 \wedge x_3) \wedge x_2' = (x_1 \vee x_3) \wedge x_3' = ((x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3))$$

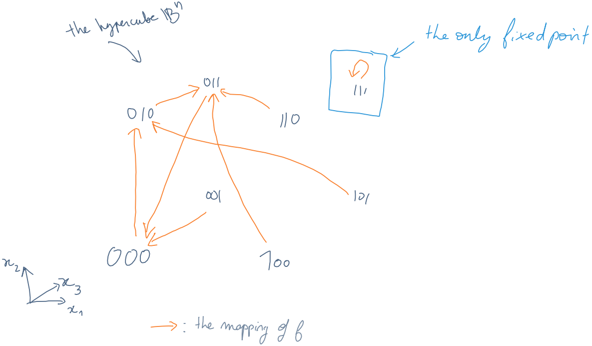


The dynamics of a Boolean network using a SAT solveur

A BN f is a function $\mathbb{B}^n \to \mathbb{B}^n$ usually expressed in propositional logics.
Transition graph: $G = (\mathbb{B}^n, E \subseteq \mathbb{B}^n \times \mathbb{B}^n)$. Is $(x, x') \in G$?
Fixpoint: a configuration $x \in \mathbb{B}^n$ such that $f(x) = x$.

Example

$$x'_1 = (x_1 \wedge x_2 \wedge x_3) \wedge x'_2 = (x_1 \vee x_3) \wedge x'_3 = ((x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3))$$

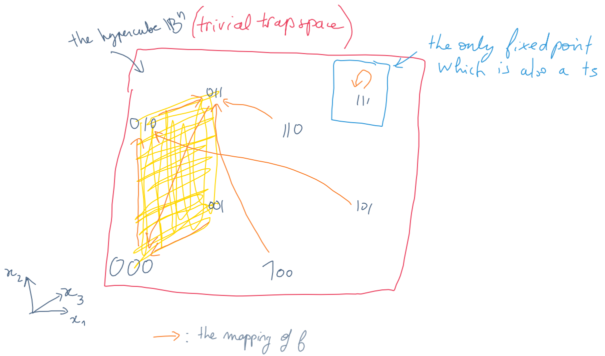


The dynamics of a Boolean network using a SAT solveur

A BN f is a function $\mathbb{B}^n \to \mathbb{B}^n$ usually expressed in propositional logics.
Transition graph: $G = (\mathbb{B}^n, E \subseteq \mathbb{B}^n \times \mathbb{B}^n)$. Is $(x, x') \in G$?
Fixpoint: a configuration $x \in \mathbb{B}^n$ such that $f(x) = x$.
Trapspace: a subcube t of $\mathbb{B}^n$ such that $\forall x \in t : f(x) \in t$

Example

$$x_1' = (x_1 \wedge x_2 \wedge x_3) \wedge x_2' = (x_1 \vee x_3) \wedge x_3' = ((x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3))$$



Boolean networks as concurent systems

The impact of updates

A Boolean network is a function $f : \mathbb{B}^n \to \mathbb{B}^n$.
Alternatively, f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).

Boolean networks as concurent systems

The impact of updates

A Boolean network is a function $f : \mathbb{B}^n \to \mathbb{B}^n$.
Alternatively, f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).

The update mode dictates which components can be updated at each time.

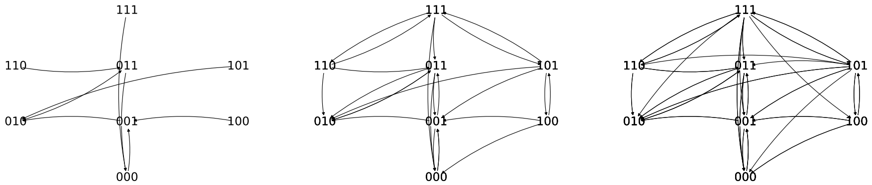
Boolean networks as concurent systems

The impact of updates

A Boolean network is a function $f : \mathbb{B}^n \to \mathbb{B}^n$.
Alternatively, f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).

The update mode dictates which components can be updated at each time.

synchronous: $\{\{A, B, C\}\}$ async.: $\{\{A\}, \{B\}, \{C\}\}$ general async.: $\mathcal{P}(\mathcal{S}) \setminus \emptyset$



Strongly impacts the dynamics $\to$ adapt your SAT constraints accordingly

Formal methods for the synthesis of Boolean networks

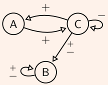
Formal synthesis of Boolean networks

Boolean network

$$\left\{ f_X : \mathbb{B}^{|\mathcal{S}|} \rightarrow \mathbb{B} \mid X \in \mathcal{S} \right\}$$

influence graph ? transition graph

Formal synthesis of Boolean networks



influences
specifications

100 is a fixpoint

$000 \longrightarrow 011 \longrightarrow 111 \longrightarrow 110$

dynamics
specifications

Boolean network

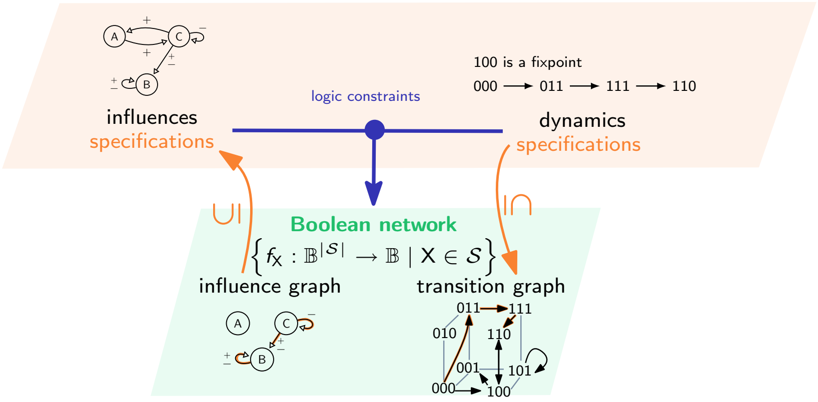
$$\left\{ f_{\mathsf{X}} : \mathbb{B}^{|\mathcal{S}|} \to \mathbb{B} \mid \mathsf{X} \in \mathcal{S} \right\}$$

influence graph

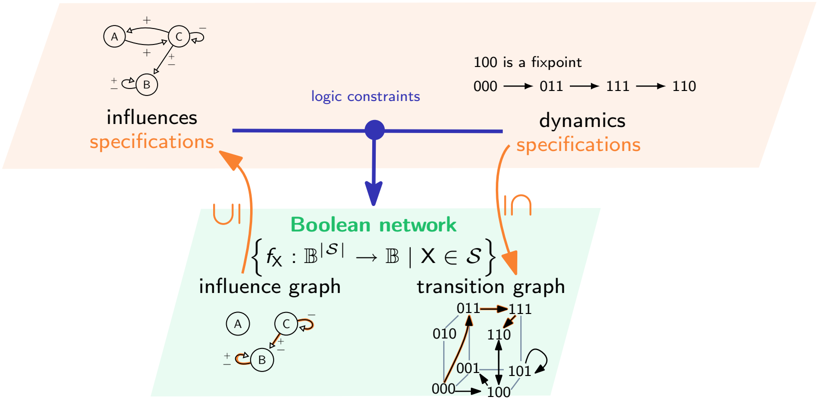
?

transition graph

Formal synthesis of Boolean networks



Formal synthesis of Boolean networks



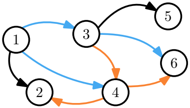
Tool: ASP (Answer set programming) provides an expressive modeling language + fast solvers

Check `BoNesis` from Loïc Paulevé and co.!

`https://bnediction.github.io/bonesis/`

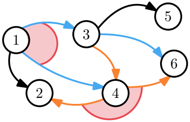
Propose fix if UNSAT

- A Boolean network f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).
- BN synthesis may be UNSAT because of conflicting dynamics specifications.



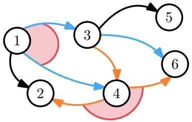
Propose fix if UNSAT

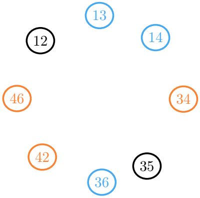
- A Boolean network f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).
- BN synthesis may be UNSAT because of conflicting dynamics specifications.



Propose fix if UNSAT

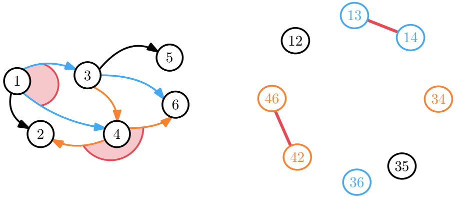
- A Boolean network f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).
- BN synthesis may be UNSAT because of conflicting dynamics specifications.
- To remove as few conflicting specifications as possible, find maximum independent sets in the conflict graph.





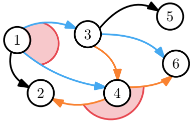
Propose fix if UNSAT

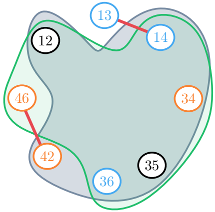
- A Boolean network f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).
- BN synthesis may be UNSAT because of conflicting dynamics specifications.
- To remove as few conflicting specifications as possible, find maximum independent sets in the conflict graph.



Propose fix if UNSAT

- A Boolean network f consists of n local functions $\mathbb{B}^n \to \mathbb{B}$ (one per species in $\mathcal{S}$).
- BN synthesis may be UNSAT because of conflicting dynamics specifications.
- To remove as few conflicting specifications as possible, find maximum independent sets in the conflict graph.





Conclusion and perspectives

To sum up

Biological system
xp / simulation
in vivo
synthesis
in silico
Models

To sum up

Biological system

xp / simulation

synthesis

in vivo

in silico

Models

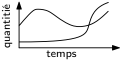
Reaction networks

$$\mathcal{R} = \left\{ \mathcal{R}_i : R_i \xrightarrow{e_i} P_i \mid i = 1 \ldots m \right\}$$

reaction, reactants, products, kinetics

$\mathcal{R} : \mathsf{A} + \mathsf{B} \xrightarrow{e} 2 * \mathsf{C}$

(mostly) time series from diff. eq.



$\rightarrow$ continuous time and values

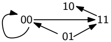
Boolean networks

$$\mathcal{B} = \left\{ f_\mathsf{X} : \mathbb{B}^{|\mathcal{S}|} \rightarrow \mathbb{B} \mid \mathsf{X} \in \mathcal{S} \right\}$$

local update function

$f_\mathsf{X} : (\mathsf{A} \wedge \mathsf{B}) \vee \neg \mathsf{C}$

Transition graph



$\rightarrow$ discrete time, boolean values

Various other formalisms

My research questions

Biological system

xp / simulation

synthesis

in vivo

in silico

Models

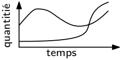
Reaction networks

$\mathcal{R} = \left\{ \mathcal{R}_i : R_i \xrightarrow{e_i} P_i \mid i = 1 \ldots m \right\}$

reaction, reactants, products, kinetics

$\mathcal{R} : A + B \xrightarrow{e} 2 * C$

(mostly) time series from diff. eq.



→ continuous time and values

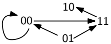
Boolean networks

$\mathcal{B} = \left\{ f_X : \mathbb{B}^{|\mathcal{S}|} \rightarrow \mathbb{B} \mid X \in \mathcal{S} \right\}$

local update function

$f_X : (A \wedge B) \vee \neg C$

Transition graph



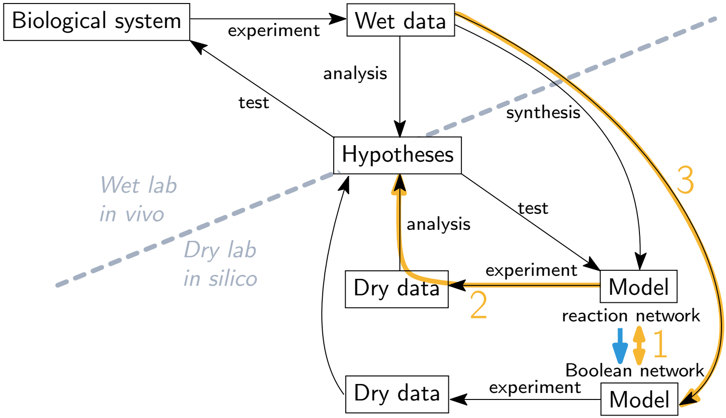
→ discrete time, boolean values

Various other formalisms

Improve synthesis and analysis of models, study abstraction relationship between the formalisms.

Thank you for your attention.

From RN to BN: the big picture



1. Formalize the relationship between RN and BN
2. Use BNs to facilitate some analyses on RN
3. Improve the BN synthesis methods

Perspectives

1. Formalize the relationship between RN and BN
   Two conjectures to investigate(*), reverse process(*)

2. Facilitate RN analyses
   Make SBML2BNET easy to use, use more evaluation criteria, include more knowledge in the synthesis, analyse FO-BNN themselves (process more RN, compute attractors(*))

3. Improve the BN synthesis methods
   Investigate, in a controled environnement
   - when we can't fullfill the constraints(*)
   - overfitting to *the* sequence of configurations?
   - impact of the choice of the binarisation procedure and error measure

Perspectives

1. Formalize the relationship between RN and BN
   Two conjectures to investigate(*), reverse process(*)

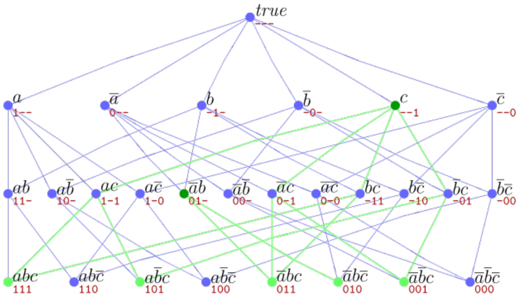
2. Facilitate RN analyses
   Make SBML2BNET easy to use, use more evaluation criteria, include more knowledge in the synthesis, **analyse FO-BNN themselves (process more RN, compute attractors(*))**

3. Improve the BN synthesis methods
   Investigate, in a controled environnement
   - **when we can't fullfill the constraints(*)**
   - overfitting to *the* sequence of configurations?
   - impact of the choice of the binarisation procedure and error measure

Minimal DNF

Given a set S of inputs for which a function f eval. to 1, each minimal-by-inclusion set of nodes that covers exactly S forms a (subset-)minimal DNF of f.
f might have several (subset-)minimal DNFs.

Example: $S = \{abc, a\bar{b}c, \bar{a}bc, \bar{a}b\bar{c}, \bar{a}\bar{b}c\}$ (light green) $\rightsquigarrow \{\bar{a}b, c\}$ (dark green)



References I

- [Cousot, Cousot, 1977]
  P. Cousot, R. Cousot,
  Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints,
  1977
- [Bornholdt, 2005]
  S. Bornholdt
  Less Is More in Modeling Large Genetic Networks,
  2005
- [Fages, Soliman, 2008a]
  F. Fages, S. Soliman,
  Abstract Interpretation and Types for Systems Biology,
  Theoretical Computer Science, vol. 403, pp. 52–70, 2008
- [Fages, Soliman, 2008b]
  F. Fages, S. Soliman,
  From Reaction Models to Influence Graphs and Back: A Theorem,
  Lecture Notes in Computer Science, pp. 90–102 2008

References II

- [Hoops et al., 2006]
  S. Hoops et al.
  COPASI—a COmplex PAthway SImulator,
  Bioinformatics, vol. 22, pp. 3067–3074 2006
- [Kohl et al., 2010]
  P. Kohl et al.
  Systems Biology: An Approach,
  Clinical Pharmacology & Therapeutics vol. 88-1 pp. 25–33 2010,
- [Lähdesmäki et al., 2003]
  H. Lähdesmäki et al.
  On Learning Gene Regulatory Networks under the Boolean Network Model,
  Machine Learning, vol. 52-1 pp. 147–167 2003,
- [Liang et al., 1998]
  S. Liang et al.
  REVEAL, a General Reverse Engineering Algorithm for Inference of Genetic Network Architectures
  Pacific Symposium on Biocomputing. pp. 18–29, 1998,

References III

- [Malik-Sheriff et al., 2020]
  R. Malik-Sheriff et al.
  BioModels—15 Years of Sharing Computational Models in Life Science
  Nucleic Acids Research vol. 48-D1, pp. D407-D415, 2020

- [Niehren et al., 2022]
  J. Niehren et al.
  Abstract Simulation of Reaction Networks via Boolean Networks
  CMSB: International Conference on Computational Methods in Systems Biology 2022,

- [Ostrowski et al., 2016]
  M. Ostrowski et al.
  Boolean Network Identification from Perturbation Time Series Data Combining Dynamics Abstraction and Logic Programming
  Biosystems vol. = 149, pp. 139–153, 2016

References IV

- [Vaginay et al., 2021]
  A. Vaginay, et al.
  Automatic Synthesis of Boolean Networks from Biological Knowledge and Data
  Communications in Computer and Information Science pp. 156–170, 2021
- [Vaginay et al., 2021]
  A. Vaginay, et al.
  From Quantitative SBML Models to Boolean Networks
  Complex Networks & Their Applications X 2021
- [Vaginay et al., 2022]
  A. Vaginay, et al.
  From Quantitative SBML Models to Boolean Networks
  Applied Network Science vol. 7-1 pp. 1–23, 2022