Synthesis of Boolean Networks from the Structure and Dynamics of Reaction Networks

Athénaïs Vaginay

7th July 2023

Rapporteurs:	François Fages	Directeur de recherche, Inria Saclay
	Loïc Paulevé	Chargé de recherche, CNRS Labri
Examinateurs:	Thierry Bastogne	Professeur Université de Lorraine
	Anna Niarakis	Maître de Conférences HDR, Université d'Evry
Encadrants:	Taha Boukhobza	Professeur Université de Lorraine
	Malika Smaïl-Tabbonne	Maître de Conférences HDR, Université de Lorraine

Systems Biology

Formal modelling and reasoning about biological systems

A set $\mathcal{S}$ of species of interest genes, proteins, cells, animals. . .

Questions

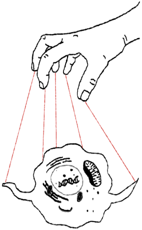
How does the system evolve?

Is the population of some cell type stable over time?

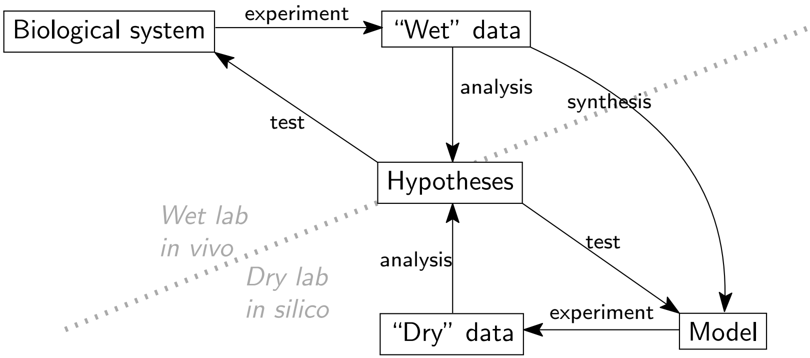
How to control the system?

Cure a pathological system

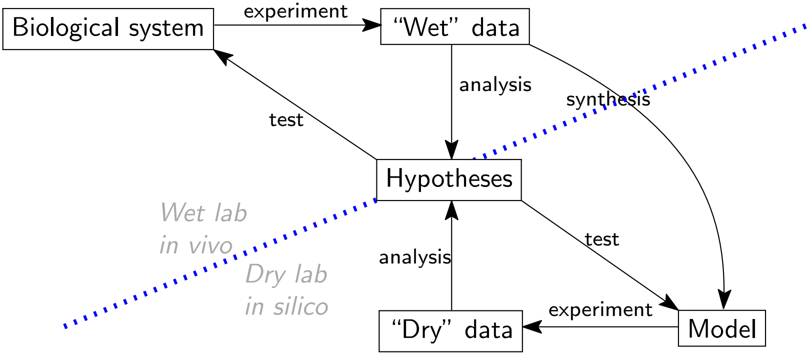
Produce more of some species of interest



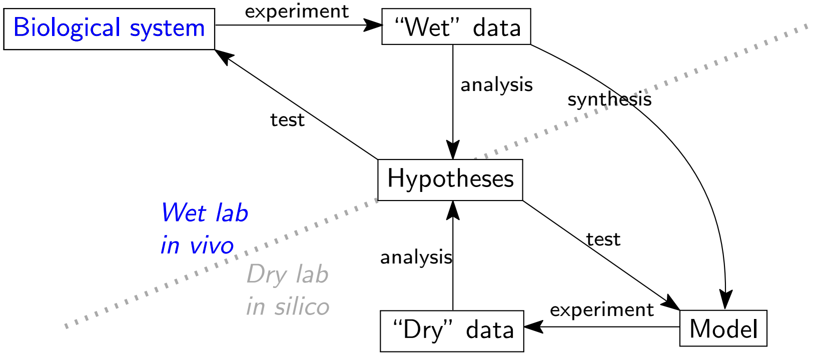
The workflow of system biology [Kohl et al., 2010]



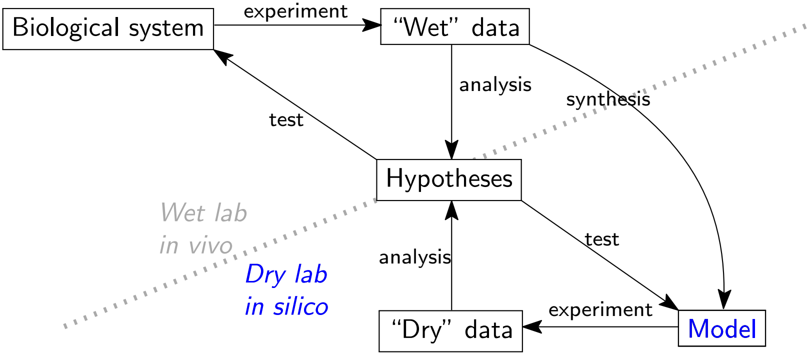
The workflow of system biology [Kohl et al., 2010]



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



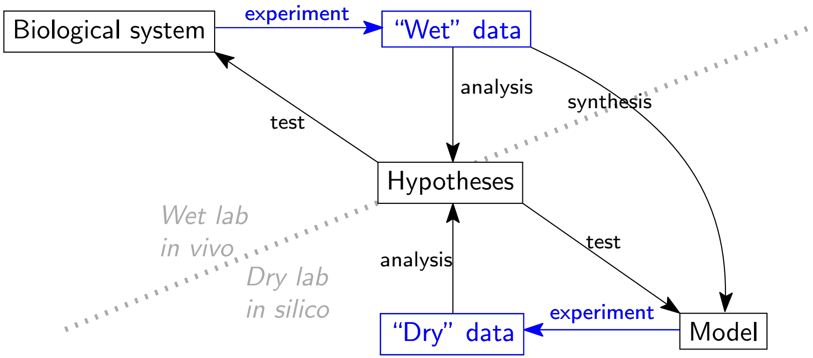
The workflow of system biology [Kohl et al., 2010]



Definition (Model)

Abstract representation (abbreviated and convenient) of the reality (more complex and detailed).

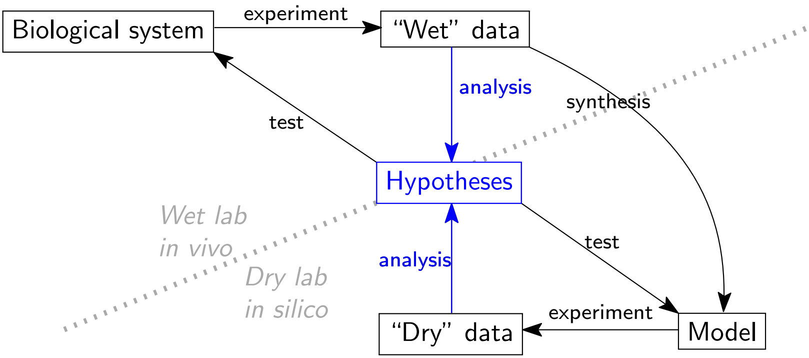
The workflow of system biology [Kohl et al., 2010]



Definition (Model)

Abstract representation (abbreviated and convenient) of the reality (more complex and detailed).

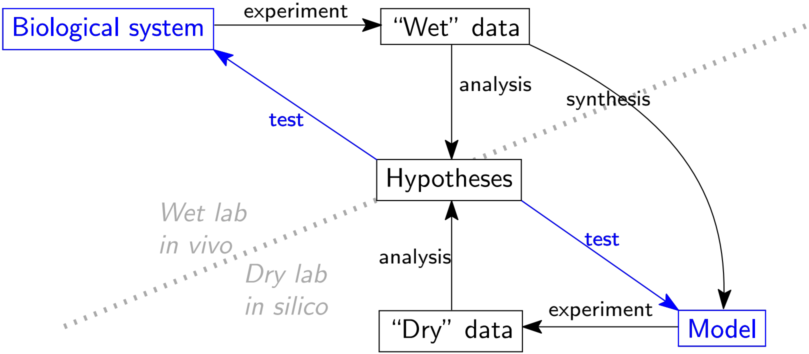
The workflow of system biology [Kohl et al., 2010]



Definition (Model)

Abstract representation (abbreviated and convenient) of the reality (more complex and detailed).

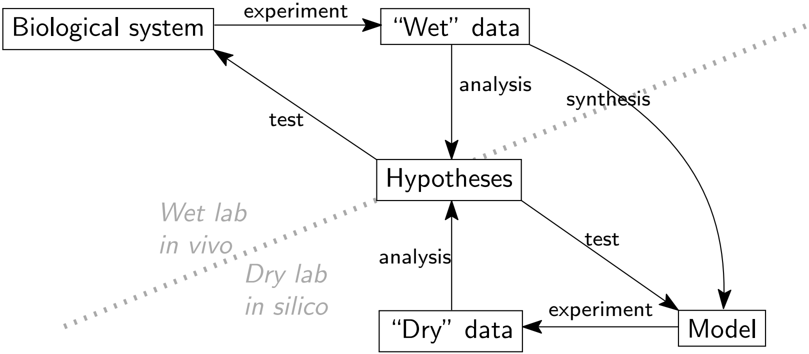
The workflow of system biology [Kohl et al., 2010]



Definition (Model)

Abstract representation (abbreviated and convenient) of the reality (more complex and detailed).

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

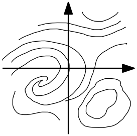


Definition (Model)

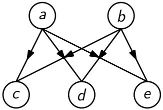
Abstract representation (abbreviated and convenient) of the reality (more complex and detailed).

A dichotomic zoo of modelling approaches

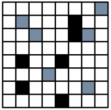
Differential equations



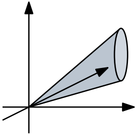
Bayesian network



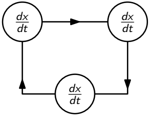
Cellular automata



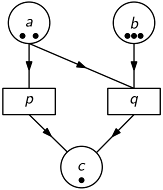
Constraint based model



Hybrid systems



Petri nets

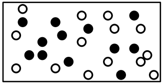


Process algebras

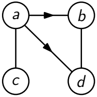
$((b(x, de)[E]) \parallel (B(y, dI)[I]))$

$bh(x, dE)bh(y, dI)(E \parallel I)$

Agent-based model



Boolean network



Reaction network

$A + B \rightarrow C$

Reaction networks, structure and dynamics

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

reaction, reactants, products, kinetics

Example

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

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

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

Reaction networks, structure and dynamics

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

reaction, reactants, products, kinetics

Example

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

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

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

Reaction networks, structure and dynamics

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

Example

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

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

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

Reaction networks, structure and dynamics

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

reaction, reactants, products, kinetics

Example

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

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

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

Reaction networks, structure and dynamics

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

Example

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

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

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

Reaction networks, structure and dynamics

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

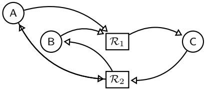
Reaction graph
$G = (\mathcal{S} \cup \mathcal{R}, E \subseteq (\mathcal{S} \times \mathcal{R}) \cup (\mathcal{R} \times \mathcal{S}))$

Example

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

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

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



Reaction networks, structure and dynamics

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

reaction, reactants, products, kinetics

Reaction graph

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

Differential semantics

ordinary differential equation (ODE)

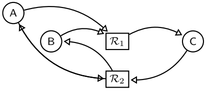
$$\left\{\dot{\mathsf{X}} = \textstyle\sum_{i\in 1\ldots m} e_i \times (R_i(\mathsf{X}) - P_i(\mathsf{X}))\right\}_{\mathsf{X}\in\mathcal{S}}$$

Example

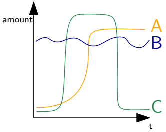
$$\mathcal{S} = \{\mathsf{A}, \mathsf{B}, \mathsf{C}\}$$

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

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



$$\begin{cases} \dot{\mathsf{A}} &= -1 \times e_1 \\ \dot{\mathsf{B}} &= \ -1 \times e_1 + 1 \times e_2 \\ \dot{\mathsf{C}} &= 2 \times e_1 + (-1) \times e_2 \end{cases}$$



Boolean network, structure and dynamics

One transition function per species in $\mathcal{S}$:
$\left\{f_X : \mathbb{B}^{|\mathcal{S}|} \to \mathbb{B}\right\}_{X \in \mathcal{S}}$ $\qquad \mathbb{B} = \{0, 1\}$

Example

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

$$f_A := 0$$
$$f_B := (B \wedge \neg C) \vee (\neg B \wedge C)$$
$$f_C := \neg C$$

Boolean network, structure and dynamics

One transition function per species in $\mathcal{S}$:
$\left\{f_{\mathsf{X}} : \mathbb{B}^{|\mathcal{S}|} \to \mathbb{B}\right\}_{\mathsf{X}\in\mathcal{S}}$ $\qquad \mathbb{B} = \{0,1\}$

Influence graph
$IG = (\mathcal{S}, E \subseteq \mathcal{S}\times\mathcal{S}, \sigma : E \to \{+,-,\pm\})$

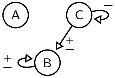
Example

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

$f_{\mathsf{A}} := 0$

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

$f_{\mathsf{C}} := \neg\mathsf{C}$



Boolean network, structure and dynamics

One transition function per species in $\mathcal{S}$:
$\left\{f_X : \mathbb{B}^{|\mathcal{S}|} \to \mathbb{B}\right\}_{X \in \mathcal{S}}$ $\quad\quad \mathbb{B} = \{0, 1\}$

Influence graph
$IG = (\mathcal{S}, E \subseteq \mathcal{S} \times \mathcal{S}, \sigma : E \to \{+, -, \pm\})$

Transition graph
$TG = (\mathbb{B}^{|\mathcal{S}|}, E \subseteq \mathbb{B}^{|\mathcal{S}|} \times \mathbb{B}^{|\mathcal{S}|})$
general asynchronous update scheme:
$\mathcal{P}(\mathcal{S}) \setminus \emptyset$

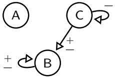
Example

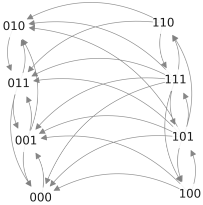
$$\mathcal{S} = \{A, B, C\}$$

$f_A := 0$

$f_B := (B \wedge \neg C) \vee (\neg B \wedge C)$

$f_C := \neg C$





Principles shared across modelling approaches

Synthesis

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

Usage

- various analyses simulation, control
- encodes our knowledge, cannot be exact

Principles shared across modelling approaches

Synthesis

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

Usage

- various analyses
  simulation, control
- encodes our knowledge, cannot be exact

Use the simplest model that contains enough information to answer the question at hand. [Bornholdt, 2005]

Principles shared across modelling approaches

Synthesis

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

Usage

- various analyses
  simulation, control
- encodes our knowledge, cannot be exact

Use the simplest model that contains enough information to answer the question at hand. [Bornholdt, 2005]
Boolean networks are simpler than reaction networks.

Principles shared across modelling approaches

Synthesis

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

Usage

- various analyses simulation, control
- encodes our knowledge, cannot be exact

Use the simplest model that contains enough information to answer the question at hand. [Bornholdt, 2005]
Boolean networks are simpler than reaction networks.

Problem statement

Automatic transformation (abstraction) of reaction networks to Boolean networks

From reactions to Boolean influences

Intuition

Input

$$\mathcal{R} : \mathrm{S} + \mathrm{E} \xrightarrow{e} \mathrm{E} + 2 \times \mathrm{P}$$

Output

$$f_{\mathrm{P}} : \mathbb{B}^{k} \to \mathbb{B}$$

input E S	output P
00	0
01	0
10	0
11	1

From reactions to Boolean influences

Intuition

Input

$$\mathcal{R} : \mathsf{S} + \mathsf{E} \xrightarrow{e} \mathsf{E} + 2 \times \mathsf{P}$$

Output

$$f_{\mathsf{P}} : \mathbb{B}^k \to \mathbb{B}$$

input	output
E S	P
00	0
01	0
10	0
11	1

Can we do this automatically? With guarantees?

From reactions to Boolean influences with guarantees

Why?

Because we want to. . .

1. Use our BNs to facilitate some analyses
2. Explore the formal relationship between RN and BN
3. Improve the BN synthesis methods

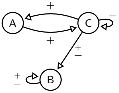
From reactions to Boolean influences with guarantees

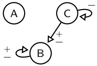
Which ones?

Input: A reaction network **Output:** A set of Boolean networks

Structure guaranty: conserve direct influences among species

IG of input RN $\supseteq$ IG of output BN



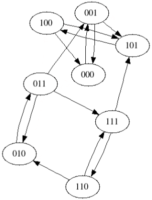


Dynamics guaranty: recover the Boolean transitions

Boolean transitions from input RN $\subseteq$ gen. async. TG of output BN

$$\left\{\dot{X} = \sum_{i\in 1\dots m} e_i \times (R_i(X) - P_i(X))\right\}_{X\in\mathcal{S}}$$

$$010 \rightarrow 011 \rightarrow 111 \rightarrow 110$$



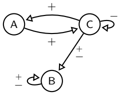
From reactions to Boolean influences with guarantees

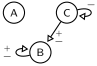
Which ones?

Input: A reaction network **Output:** A set of Boolean networks

Structure guaranty: conserve direct influences among species

IG of input RN ⊇ IG of output BN



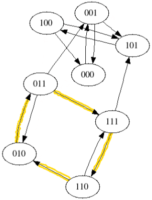


Dynamics guaranty: recover the Boolean transitions

Boolean transitions from input RN ⊆ gen. async. TG of output BN

$$\left\{\dot{\mathsf{X}} = \textstyle\sum_{i\in 1\ldots m} e_i \times (R_i(\mathsf{X}) - P_i(\mathsf{X}))\right\}_{\mathsf{X}\in\mathcal{S}}$$

$010 \rightarrow 011 \rightarrow 111 \rightarrow 110$



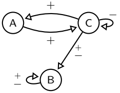
From reactions to Boolean influences with guarantees

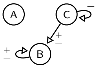
Which ones?

Input: A reaction network **Output:** A set of Boolean networks

Structure guaranty: conserve direct influences among species

IG of input RN ⊇ IG of output BN



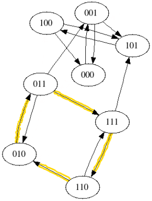


Dynamics guaranty: recover the Boolean transitions

Boolean transitions from input RN ⊆ gen. async. TG of output BN

$$\left\{\dot{X} = \sum\nolimits_{i\in 1\dots m} e_i \times (R_i(X) - P_i(X))\right\}_{X\in\mathcal{S}}$$

010 → 011 → 111 → 110



Coverage: proportion of recovered transitions (ideally 100%)

Outline

1. My method, its implementation and guarantees

SBML2BNET

STEP 1: Retrieve from the input reaction network

Structure: influence graph
Dynamics: Boolean transitions

- binarised time series from classic simulation of the ODEs
- abstract simulation of the ODEs

STEP 2: BN synthesis with `ASK&D-BN`

2. Evaluation of the approach

3. Link to other abstractions

4. Conclusion and perspectives

Outline

1. My method, its implementation and guarantees

SBML2BNET

STEP 1: Retrieve from the input reaction network

Structure: influence graph

Dynamics: Boolean transitions

- binarised time series from classic simulation of the ODEs
- abstract simulation of the ODEs

STEP 2: BN synthesis with `ASK&D-BN`

2. Evaluation of the approach

3. Link to other abstractions

4. Conclusion and perspectives

STEP 1: Retrieve an influence graph and Boolean transitions

Running example $\mathcal{R}_{\text{enz}}$

$$\underbrace{\overbrace{\mathsf{E} + \mathsf{S} \xleftrightarrow[e_{\mathrm{off}}]{e_{\mathrm{on}}} \mathsf{C}}^{r_{\mathrm{on}}}}_{r_{\mathrm{off}}} \overbrace{\xrightarrow{e_{\mathrm{cat}}} \mathsf{E} + 2 \times \mathsf{P}}^{r_{\mathrm{cat}}}$$

Its ODEs

$$\begin{cases} \dot{\mathsf{C}} &= (e_{\mathrm{on}} - e_{\mathrm{off}}) - e_{\mathrm{cat}} \\ \dot{\mathsf{P}} &= 2 \times e_{\mathrm{cat}} \\ \dot{\mathsf{S}} &= -(e_{\mathrm{on}} - e_{\mathrm{off}}) \\ \dot{\mathsf{E}} &= -(e_{\mathrm{on}} - e_{\mathrm{off}}) + e_{\mathrm{cat}} \end{cases}$$

Its parameters

$$\begin{aligned} e_{\mathrm{on}} &= 10^6 \times \mathsf{E} \times \mathsf{S} \\ e_{\mathrm{off}} &= 0.2 \times \mathsf{C} \\ e_{\mathrm{cat}} &= 0.1 \times \mathsf{C} \end{aligned}$$

Retrieve the influence graph of a reaction network

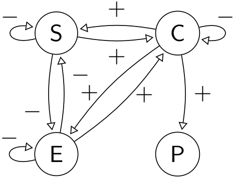
Contribution

Implement the routines from [Fages, Soliman, 2008b]

Guarantees

Overapproximates the possible signs of $\frac{\partial \dot{X}}{\partial Y}$
→ capture all the direct influences between the species

Influence graph of $\mathcal{R}_{\text{enz}}$



Retrieve Boolean transitions from a reaction network

Numerical simulation and binarisation

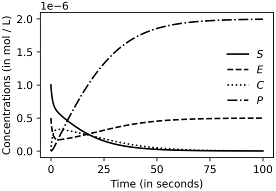
Contribution

Use dedicated tools for simulation
Apply binarisation procedure

Guarantees

Approximate the real solution of the ODE with good accuracy [Hoops et al., 2006]
but loose causation

For $\mathcal{R}_{\text{enz}}$



Binarisation	Boolean configuration sequence
Midrange	1100 → 1000 → 1010 → 0010 → 0011 → 0101
Median	1100 → 1010 → 0011 → 0101
Mean	1100 → 1010 → 1000 → 0011 → 0101
Above 0	1100 → 1111 → 1011 → 1111 → 0111
Expected	1100 → $**10$ → $***1$

Retrieve Boolean transitions from a reaction network

Abstract simulation — Intuition

Joint work with Joachim Niehren and Cristian Versari [Niehren et al., 2022]

Use the rule of signs to reason on the causal relationship between the signs of the variables values of the ODE system species amount and derivatives

X is above 0 if...

- it was already above 0 and its derivative was null (no change) $\textit{plus} + 0 = \textit{plus}$
- it was already above 0 and its derivative was positive (X has increased) $\textit{plus} + \textit{plus} = \textit{plus}$

If X was above 0 and its derivative was negative (X has decreased) $\textit{plus} - \textit{plus} = \textit{unknown} \rightsquigarrow$ nondeterminism

Retrieve Boolean transitions from a reaction network

Abstract simulation — In practice

Contribution

$$\mathcal{V} = \bigcup_{\mathsf{X} \in \mathcal{S}} \{\mathsf{X}, \mathring{\mathsf{X}}, \underset{\text{next}}{\mathsf{X}}, \underset{\text{next}}{\mathring{\mathsf{X}}}\}$$

Causal relationships encoded by a first-order logic formula ϕ

Solve ϕ on the structure of signs $\mathbb{S} = \{-1, 0, 1\}$

Restrict the solutions on $\mathcal{S} \rightsquigarrow$ relation $\mathbb{B}^{|\mathcal{S}|} \times \mathbb{B}^{|\mathcal{S}|}$

FOBNN: First-Order Boolean networks with nondeterministic updates

Guarantee

- Keep the causalities of changes
- Proof of correctness: overapproximation of an ideal Euler simulation

Retrieve Boolean transitions from a reaction network

Abstract simulation — Example on $\mathcal{R}_{\mathsf{enz}}$

$$
\begin{aligned}
& \mathring{S} = -e_{\mathrm{on}} + e_{\mathrm{off}} && \wedge\ \underset{\mathrm{next}}{\mathring{S}} = -\underset{\mathrm{next}}{e_{\mathrm{on}}} + \underset{\mathrm{next}}{e_{\mathrm{off}}} \\
\wedge\ & \mathring{E} = -e_{\mathrm{on}} + e_{\mathrm{off}} + e_{\mathrm{cat}} && \wedge\ \underset{\mathrm{next}}{\mathring{E}} = -\underset{\mathrm{next}}{e_{\mathrm{on}}} + \underset{\mathrm{next}}{e_{\mathrm{off}}} + \underset{\mathrm{next}}{e_{\mathrm{cat}}} \\
\wedge\ & \mathring{C} = e_{\mathrm{on}} - e_{\mathrm{off}} - e_{\mathrm{cat}} && \wedge\ \underset{\mathrm{next}}{\mathring{C}} = \underset{\mathrm{next}}{e_{\mathrm{on}}} - \underset{\mathrm{next}}{e_{\mathrm{off}}} - \underset{\mathrm{next}}{e_{\mathrm{cat}}} \\
\wedge\ & \mathring{P} = e_{\mathrm{cat}} && \wedge\ \underset{\mathrm{next}}{\mathring{P}} = \underset{\mathrm{next}}{e_{\mathrm{cat}}}
\end{aligned}
$$

$$
\begin{aligned}
\wedge\ & \underset{\mathrm{next}}{S} = S + \mathring{S} && \wedge\ S \leq \underset{\mathrm{next}}{S} \\
\wedge\ & \underset{\mathrm{next}}{E} = E + \mathring{E} && \wedge\ E \leq \underset{\mathrm{next}}{E} \\
\wedge\ & \underset{\mathrm{next}}{C} = C + \mathring{C} && \wedge\ C \leq \underset{\mathrm{next}}{C} \\
\wedge\ & \underset{\mathrm{next}}{P} = P + \mathring{P} && \wedge\ P \leq \underset{\mathrm{next}}{P}
\end{aligned}
$$

with

$$
\begin{aligned}
& e_{\mathrm{on}} = 10^6 \times S \times E && e_{\mathrm{off}} = 0.2 \times C && e_{\mathrm{cat}} = 0.1 \times C \\
& \underset{\mathrm{next}}{e_{\mathrm{on}}} = 10^6 \times \underset{\mathrm{next}}{S} \times \underset{\mathrm{next}}{E} && \underset{\mathrm{next}}{e_{\mathrm{off}}} = 0.2 \times \underset{\mathrm{next}}{C} && \underset{\mathrm{next}}{e_{\mathrm{cat}}} = 0.1 \times \underset{\mathrm{next}}{C}
\end{aligned}
$$

Retrieve Boolean transitions from a reaction network

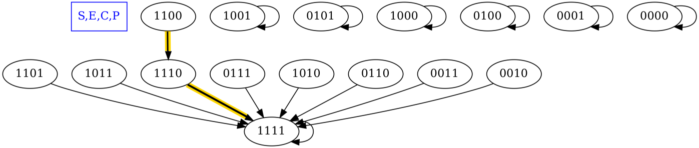
Abstract simulation — Result on $\mathcal{R}_{\text{enz}}$

Application to $\mathcal{R}_{\text{enz}}$

$$\mathcal{R}_{\text{on}} : \mathsf{S} + \mathsf{E} \xrightarrow{\mathrm{k_{on}} \times \mathsf{S} \times \mathsf{E}} \mathsf{C}$$

$$\mathcal{R}_{\text{off}} : \mathsf{C} \xrightarrow{\mathrm{k_{off}} \times \mathsf{C}} \mathsf{S} + \mathsf{E}$$

$$\mathcal{R}_{\text{cat}} : \mathsf{C} \xrightarrow{\mathrm{k_{cat}} \times \mathsf{C}} \mathsf{E} + 2 \times \mathsf{P}$$



Summary about our guarantees

Structure: Influence graph
from syntactic parsing of the reactions

- captures all the direct influences among species

Dynamics: Boolean transitions
from numerical simulation of the ODEs + binarisation

- good approximation or the analytical solution
- but we lose causality

from abstract simulation of the ODEs

- correct overapproximation of perfect Euler that captures causality

Outline

1. My method, its implementation and guarantees

SBML2BNET

STEP 1: Retrieve from the input reaction network

Structure: influence graph

Dynamics: Boolean transitions

- binarised time series from classic simulation of the ODEs
- abstract simulation of the ODEs

STEP 2: BN synthesis with `ASK&D-BN`

2. Evaluation of the approach

3. Link to other abstractions

4. Conclusion and perspectives

STEP 2:
Boolean network synthesis with `ASK&D-BN`

ASK&D-BN [Vaginay et al., 2021]

Input	Output
Structure Influence graph **Dynamics** Time series / Boolean time series List of Boolean transitions	Set of *compatible* Boolean networks

1. **Local search** species-wise synthesis of *all* the transition functions compatible with the given influence graph and time series

2. **Global assembly** produce all the possible BNs

ASK&D-BN [Vaginay et al., 2021]

Input	Output
Structure Influence graph **Dynamics** Time series / Boolean time series List of Boolean transitions	Set of *compatible* Boolean networks

1. **Local search** species-wise synthesis of *all* the transition functions compatible with the given influence graph and time series
   Generate candidates → Structure constraint → Dynamic constraint → Minimality constraint

2. **Global assembly** produce all the possible BNs

ASK&D-BN [Vaginay et al., 2021]

Input	Output
Structure Influence graph **Dynamics** Time series / Boolean time series List of Boolean transitions	Set of *compatible* Boolean networks

1. **Local search** species-wise synthesis of *all* the transition functions compatible with the given influence graph and time series
   Generate candidates → Structure constraint → Dynamic constraint → Minimality constraint
   ⇝ Answer-Set Programming

   ```
   head :- condition1; condition2.
   ```

2. **Global assembly** produce all the possible BNs

ASK&D-BN— Local search

Candidate transition function

Search space: $2^{3^{|\mathcal{S}|}}$ non-redundant DNF = non-redundant disjunction of non-redundant conjunctions

ASK&D-BN— Local search

Candidate transition function

Search space: $2^{3^{|\mathcal{S}|}}$ non-redundant DNF = non-redundant disjunction of non-redundant conjunctions

Pick a subset of non-redundant conjunctions

```
% GIVEN non-redundant conjunctions 𝒮 → {0,1,−1}
% conj(ID, Species, Sign}
conj(1, a,  1). conj(1, b,-1). conj(1, c,  0). %  A ∧ ¬B
conj(2, a, -1). conj(2, b, 0). conj(2, c, -1). % ¬A ∧ ¬C
conj(3, a, -1). conj(3, b,-1). conj(3, c, -1). % ¬A ∧ ¬B ∧ ¬C
...
1{conjTakenID(0..maxNbPossibleConj)}. % choice rule
```

ASK&D-BN— Local search

Candidate transition function

Search space: $2^{3^{|\mathcal{S}|}}$ non-redundant DNF = non-redundant disjunction of non-redundant conjunctions

Pick a subset of non-redundant conjunctions

```
% GIVEN non-redundant conjunctions 𝒮 → {0, 1, −1}
% conj(ID, Species, Sign}
conj(1, a,  1). conj(1, b,-1). conj(1, c,  0). %  A ∧ ¬B
conj(2, a, -1). conj(2, b, 0). conj(2, c, -1). % ¬A ∧ ¬C
conj(3, a, -1). conj(3, b,-1). conj(3, c, -1). % ¬A ∧ ¬B ∧ ¬C
...
1{conjTakenID(0..maxNbPossibleConj)}. % choice rule
```

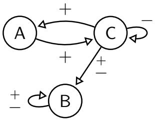
Example

`conjTakenID(1). conjTakenID(2).` ⇒ candidate = $(A \land \neg B) \lor (\neg A \land \neg C)$

ASK&D-BN— Local search

Structure constraints

influence graph of the Boolean network ⊆ influence graph of the reaction network

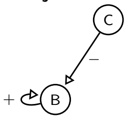


Do not select a conjunction that uses a forbidden literal

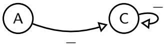
```
% conjTaken(ID, Species, Sign}
ig(ParentID, x, V) :- conjTaken(ConjID, ParentID, V); V!=0.
:- ig(ParentID, x, V) ; not pig(ParentID, x, V).
```

Example

valid conjunction: ¬C ∧ B



invalid conjunction: ¬A ∧ ¬C



ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

putative input	output

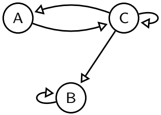
ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

input influence graph (unsigned)



putative input	output
C	A
BC	B
AC	C

ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

$$010 \xrightarrow[\textcircled{1}]{} 011 \xrightarrow[\textcircled{2}]{} 100 \xrightarrow[\textcircled{3}]{} 001$$

putative input	output
C	A
BC	B
AC	C

ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

$$010 \xrightarrow[\textcircled{1}]{C} 011 \xrightarrow[\textcircled{2}]{A,B,C} 100 \xrightarrow[\textcircled{3}]{A,C} 001$$

Do not assume the underlying update scheme

putative input	output
C	A
BC	B
AC	C

ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

$$010 \xrightarrow[\textcircled{1}]{\text{C}} 011 \xrightarrow[\textcircled{2}]{\text{A,B,C}} 100 \xrightarrow[\textcircled{3}]{\text{A,C}} 001$$

Do not assume the underlying update scheme

putative input	output	
C	A	
1	1	$\textcircled{2}$
BC	B	
AC	C	

ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

$$010 \xrightarrow[①]{C} 011 \xrightarrow[②]{A,B,C} 100 \xrightarrow[③]{A,C} 001$$

Do not assume the underlying update scheme

putative input	output	
C	A	
0	0	③
1	1	②
BC	B	
AC	C	

ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

$$010 \xrightarrow[\textcircled{1}]{C} 011 \xrightarrow[\textcircled{2}]{A,B,C} 100 \xrightarrow[\textcircled{3}]{A,C} 001$$

Do not assume the underlying update scheme

putative input	output	
C	A	
0	0	③
1	1	②
BC	B	
11	0	②
AC	C	
00	1	①
01	0	②
10	1	③

ASK&D-BN— Local search

Dynamics constraints

(1) Build partial truth tables for each species X

What were the values of the putative inputs of species X when its value changed?

$$010 \xrightarrow[\textcircled{1}]{C} 011 \xrightarrow[\textcircled{2}]{A,B,C} 100 \xrightarrow[\textcircled{3}]{A,C} 001$$

Do not assume the underlying update scheme

putative input	output	
C	A	
0	0	③
1	1	②
BC	B	
11	0	②
AC	C	
00	1	①
01	0	②
10	1	③

Truth table of a candidate function vs reconstructed truth tables

ASK&D-BN— Local search

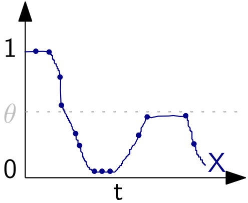
Dynamics constraints

(2) Minimise the error (ideally 0)

```
#minimize{E@2 : error(E)}. % highest priority
```

X_t: continuous value of X at time t

θ: binarisation threshold for X



ASK&D-BN— Local search

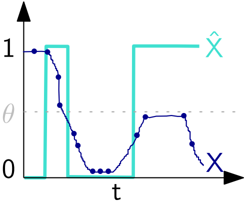
Dynamics constraints

(2) Minimise the error (ideally 0)

```
#minimize{E@2 : error(E)}. % highest priority
```

X_t: continuous value of X at time t

θ: binarisation threshold for X



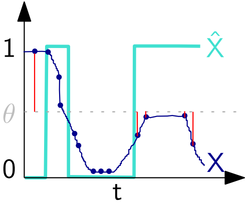
ASK&D-BN— Local search

Dynamics constraints

(2) Minimise the error (ideally 0)

```
#minimize{E@2 : error(E)}. % highest priority
```

X_t: continuous value of X at time t
θ: binarisation threshold for X
$\mathcal{U}$: set of unexplained time steps



ASK&D-BN— Local search

Dynamics constraints

(2) Minimise the error (ideally 0)

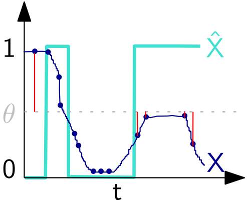
```
#minimize{E@2 : error(E)}. % highest priority
```

X_t: continuous value of X at time t

θ: binarisation threshold for X

$\mathcal{U}$: set of unexplained time steps

$\mathtt{E} = \sum_{t \in \mathcal{U}} |\theta - X_t|$



ASK&D-BN— Local search

Minimality constraint

Select the smallest expressions

Minimality → most general conditions

```
sizeconj(C, S):-conjTakenID(C);S=#sum{|V|,N:conj(C, N, V)} .
sizeDNF(S):- S=#sum{N,C: sizeconj(C, N), conjTakenID(C)} .
% N elements in conjunction C
#minimize{S@1 : sizeDNF(S)}. % lower priority
```

putative input	observed output
AB	X
00	
01	0
10	1
11	

ASK&D-BN— Local search

Minimality constraint

Select the smallest expressions

Minimality → most general conditions

```
sizeconj(C, S):-conjTakenID(C);S=#sum{|V|,N:conj(C, N, V)} .
sizeDNF(S):- S=#sum{N,C: sizeconj(C, N), conjTakenID(C)} .
% N elements in conjunction C
#minimize{S@1 : sizeDNF(S)}. % lower priority
```

putative input	observed output	possible completions			
AB	X				
00		0	1	0	1
01	0	0	0	0	0
10	1	1	1	1	1
11		0	0	1	1

ASK&D-BN— Local search

Minimality constraint

Select the smallest expressions

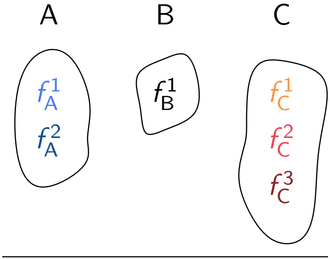
Minimality → most general conditions

```
sizeconj(C, S):-conjTakenID(C);S=#sum{|V|,N:conj(C, N, V)} .
sizeDNF(S):- S=#sum{N,C: sizeconj(C, N), conjTakenID(C)} .
% N elements in conjunction C
#minimize{S@1 : sizeDNF(S)}. % lower priority
```

putative input	observed output	possible completions			
AB	X				
00		0	1	0	1
01	0	0	0	0	0
10	1	1	1	1	1
11		0	0	1	1
	smallest expressions	$A \land \neg B$	$\neg B$	A	$A \lor \neg B$
	size:	2	1	1	2

ASK&D-BN— Global assembly

Cartesian product of the set of transition functions synthesised for each species



A B C

f_A^1 f_A^2

f_B^1

f_C^1 f_C^2 f_C^3

$$\mathcal{B}_1 = \{ f_A^1, f_B^1, f_C^1 \}$$

$$\mathcal{B}_2 = \{ f_A^1, f_B^1, f_C^2 \}$$

$$\mathcal{B}_3 = \{ f_A^1, f_B^1, f_C^3 \}$$

$$\mathcal{B}_4 = \{ f_A^2, f_B^1, f_C^1 \}$$

$$\mathcal{B}_5 = \{ f_A^2, f_B^1, f_C^2 \}$$

$$\mathcal{B}_6 = \{ f_A^2, f_B^1, f_C^3 \}$$

Outline

1. My method, its implementation and guarantees

SBML2BNET

STEP 1: Retrieve from the input reaction network

Structure: influence graph

Dynamics: Boolean transitions

- binarised time series from classic simulation of the ODEs
- abstract simulation of the ODEs

STEP 2: BN synthesis with `ASK&D-BN`

2. Evaluation of the approach

3. Link to other abstractions

4. Conclusion and perspectives

Evaluation of the approach

Evaluation of SBML2BNET

1. The BN synthesis itself [Vaginay et al., 2021]:
   `ASK&D-BN` versus other state-of-the-art methods

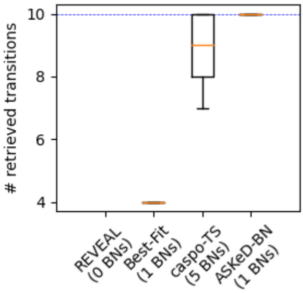
2. One specific variant of the complete approach to real-world reaction networks [Vaginay et al., 2021, Vaginay et al., 2022]:
   influence graph + time series and midrange binarisation

3. Several variants of the complete approach on $\mathcal{R}_{\text{enz}}$:
   compare concrete and abstract simulation

ASK&D-BN versus REVEAL[1], Best-Fit[2] and Caspo-TS[3]

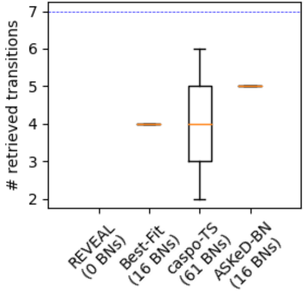
A. thaliana

5 components, 10 transitions



yeast

4 components, 7 transitions

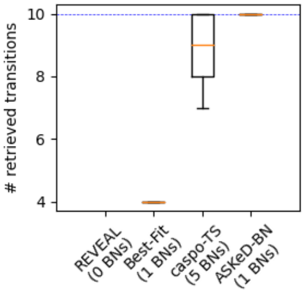


[1][Liang et al., 1998] [2][Lähdesmäki et al., 2003] [3][Ostrowski et al., 2016]

ASK&D-BN versus REVEAL[1], Best-Fit[2] and Caspo-TS[3]

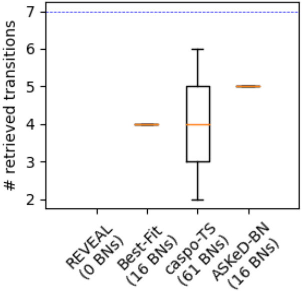
A. thaliana

5 components, 10 transitions



yeast

4 components, 7 transitions

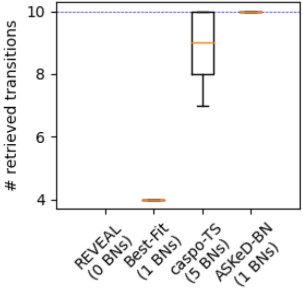


- REVEAL fails

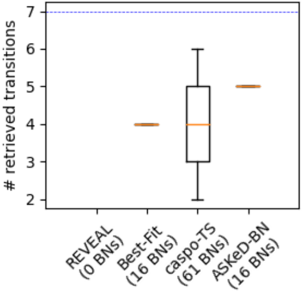
[1] [Liang et al., 1998] [2] [Lähdesmäki et al., 2003] [3] [Ostrowski et al., 2016]

ASK&D-BN versus REVEAL[1], Best-Fit[2] and Caspo-TS[3]

A. thaliana
5 components, 10 transitions



yeast
4 components, 7 transitions

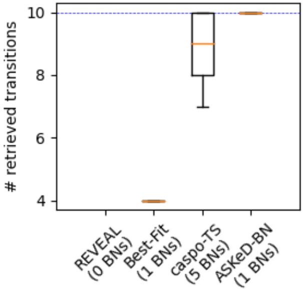


- REVEAL fails
- Best-Fit lacks consistency

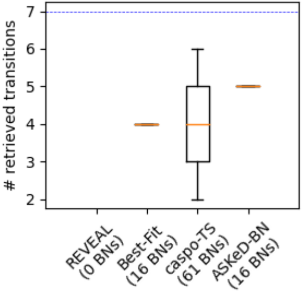
[1] [Liang et al., 1998] [2] [Lähdesmäki et al., 2003] [3] [Ostrowski et al., 2016]

ASK&D-BN versus REVEAL[1], Best-Fit[2] and Caspo-TS[3]

A. thaliana
5 components, 10 transitions



yeast
4 components, 7 transitions

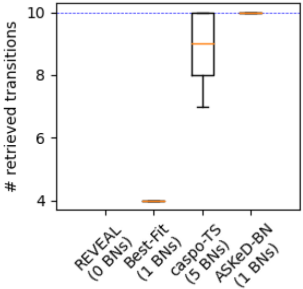


- REVEAL fails
- Best-Fit lacks consistency
- Caspo-TS returns more BNs, some of them with poor coverage because of reachability

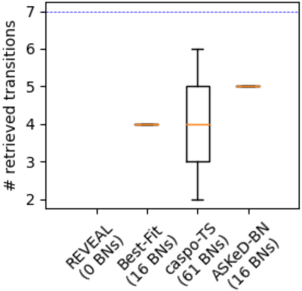
[1] [Liang et al., 1998] [2] [Lähdesmäki et al., 2003] [3] [Ostrowski et al., 2016]

ASK&D-BN versus REVEAL[1], Best-Fit[2] and Caspo-TS[3]

A. thaliana
5 components, 10 transitions



yeast
4 components, 7 transitions



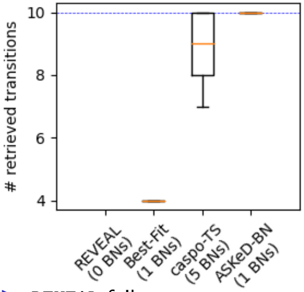
- REVEAL fails
- Best-Fit lacks consistency
- Caspo-TS returns more BNs, some of them with poor coverage because of reachability
- ASK&D-BN returns a small number of BN, with good coverage and low variance ✓

[1] [Liang et al., 1998] [2] [Lähdesmäki et al., 2003] [3] [Ostrowski et al., 2016]

ASK&D-BN versus REVEAL[1], Best-Fit[2] and Caspo-TS[3]

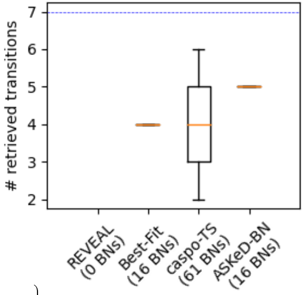
A. thaliana

5 components, 10 transitions



yeast

4 components, 7 transitions



- REVEAL fails
- Best-Fit lacks consistency
- Caspo-TS returns more BNs, some of them with poor coverage because of reachability
- ASK&D-BN returns a small number of BN, with good coverage and low variance ✓

confirmed on $>$ 300 datasets
generated from existing BNs
from the repository of PyBoolNet

[1] [Liang et al., 1998] [2] [Lähdesmäki et al., 2003] [3] [Ostrowski et al., 2016]

Results to real-world reaction networks (from BioModels[4])

Input: an **extended** reaction network rules and events
Output: a set of compatible Boolean networks, according to `ASK&D-BN`

SBML2BNET variant:

1. structure constraint (**extended** influence graph)
2. soft dynamics constraints (time series and midrange binarisation)
3. `ASK&D-BN`

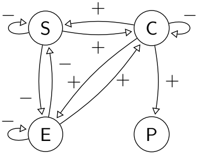
Result:

- on 155 reaction networks processed in less than 30 hours
- we synthesise perfect Boolean networks for ∼90% of them ✓
  139/155 sets of BNs have a coverage proportion median = 1

[4] [Malik-Sheriff et al., 2020]

A closer look: $\mathcal{R}_{\mathrm{enz}}$

$$\underbrace{\overbrace{\mathrm{E}+\mathrm{S} \xrightleftharpoons[e_{\mathrm{off}}]{e_{\mathrm{on}}} \mathrm{C}}^{r_{\mathrm{on}}}}_{r_{\mathrm{off}}} \overbrace{\xrightarrow{e_{\mathrm{cat}}} \mathrm{E}+2\times\mathrm{P}}^{r_{\mathrm{cat}}}$$



Setting n°1

- influence graph
- time series
- binarised time series

midrange (0.8) and median (0.6):

$$
\begin{aligned}
f_{\mathrm{S}} &:= \neg \mathrm{E}\\
f_{\mathrm{E}} &:= \neg \mathrm{S}\\
f_{\mathrm{C}} &:= \mathrm{S}\\
f_{\mathrm{P}} &:= \mathrm{C}
\end{aligned}
$$

→ Coverage depends on the binarisation procedure, BNs miss some influences

Setting n°2

- full graph from abstract simulation

$$
\begin{aligned}
f_{\mathrm{S}} &:= \mathrm{C} \vee \mathrm{S}\\
f_{\mathrm{E}} &:= \mathrm{E} \vee \mathrm{C}\\
f_{\mathrm{C}} &:= (\mathrm{E} \wedge \mathrm{C}) \vee \mathrm{C}\\
f_{\mathrm{P}} &:= \mathrm{C} \vee \mathrm{P}
\end{aligned}
$$

→ Perfect coverage but does not comply with the influence graph

$\Longrightarrow$ They do not capture the same thing

Outline

1. My method, its implementation and guarantees

SBML2BNET

STEP 1: Retrieve from the input reaction network

Structure: influence graph

Dynamics: Boolean transitions

- binarised time series from classic simulation of the ODEs
- abstract simulation of the ODEs

STEP 2: BN synthesis with `ASK&D-BN`

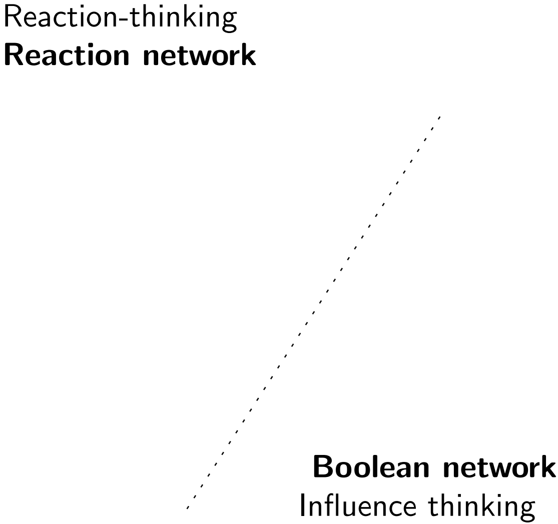
2. Evaluation of the approach

3. Link to other abstractions

4. Conclusion and perspectives

Relation to other abstractions

Our abstraction versus other abstractions



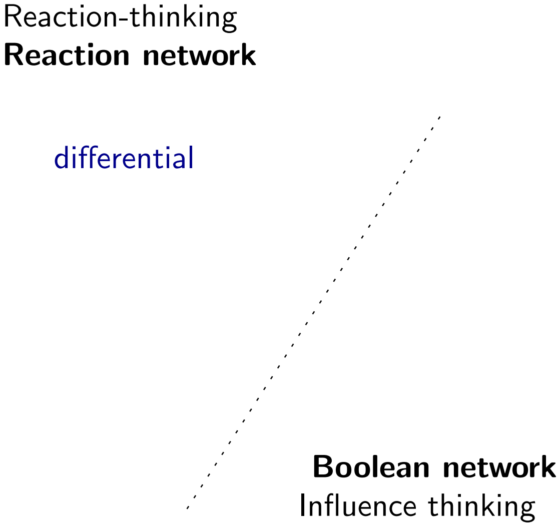
Reaction-thinking

Reaction network

Boolean network

Influence thinking

Our abstraction versus other abstractions



Reaction-thinking

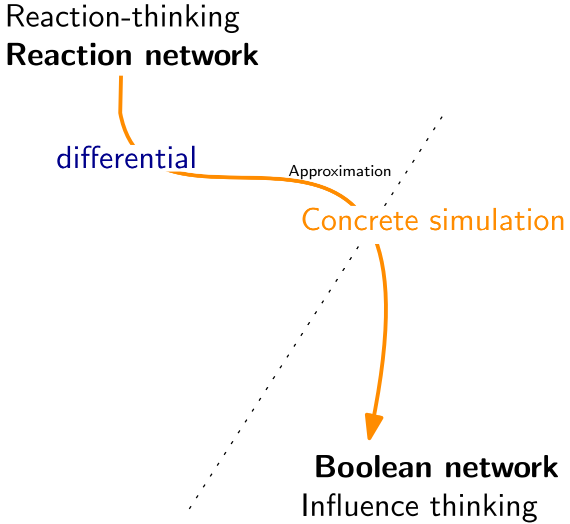
Reaction network

differential

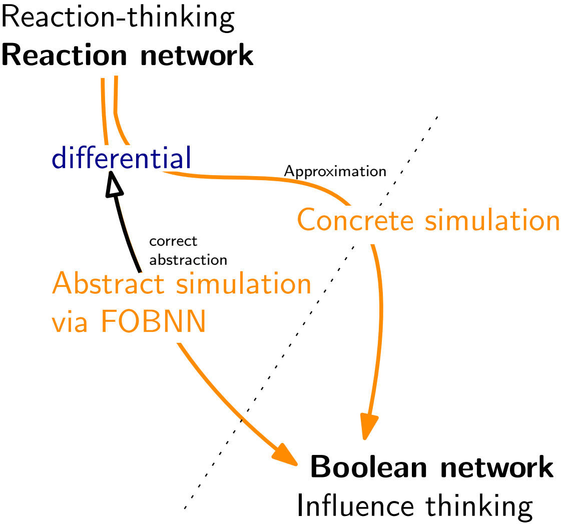
Boolean network

Influence thinking

Our abstraction versus other abstractions



Our abstraction versus other abstractions



Reaction-thinking

Reaction network

differential

Approximation

Concrete simulation

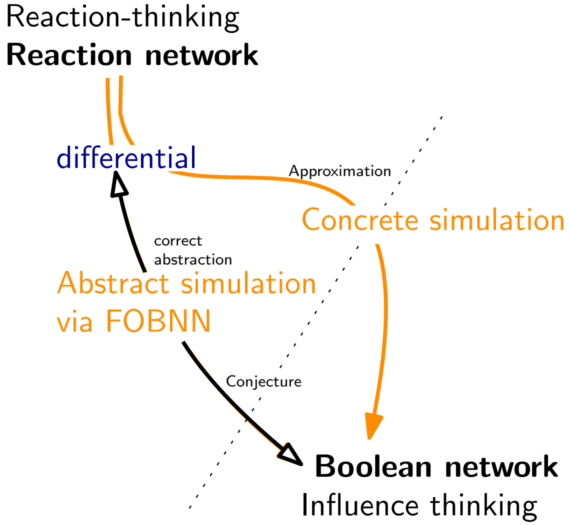
correct
abstraction

Abstract simulation
via FOBNN

Boolean network

Influence thinking

Our abstraction versus other abstractions



Reaction-thinking

Reaction network

differential

Approximation

Concrete simulation

correct abstraction

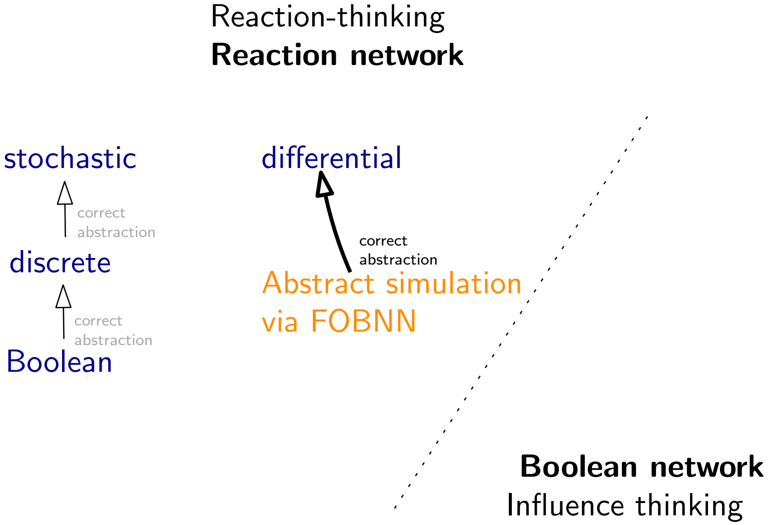
Abstract simulation via FOBNN

Conjecture

Boolean network

Influence thinking

Our abstraction versus other abstractions



Reaction-thinking

Reaction network

stochastic

correct abstraction

discrete

correct abstraction

Boolean

differential

correct abstraction

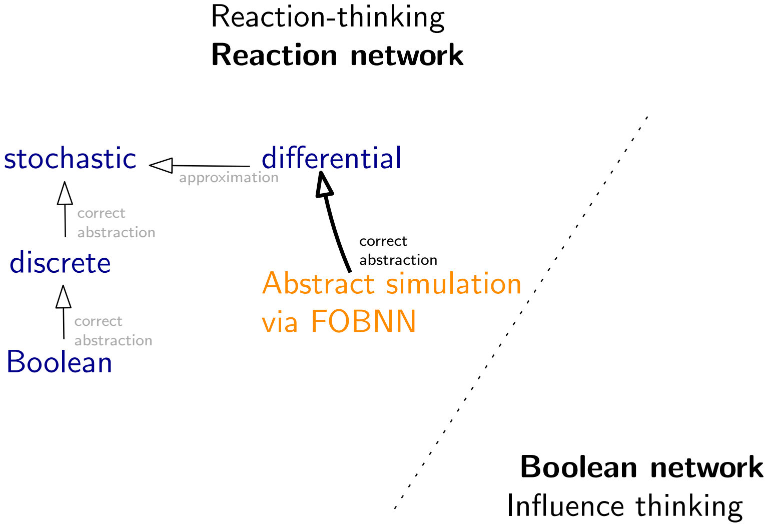
Abstract simulation via FOBNN

Boolean network

Influence thinking

[Fages, Soliman, 2008a]

Our abstraction versus other abstractions



Reaction-thinking

Reaction network

stochastic

approximation

differential

correct abstraction

discrete

correct abstraction

Boolean

correct abstraction

Abstract simulation via FOBNN

Boolean network

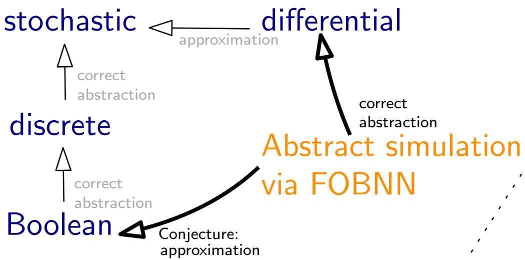
Influence thinking

[Fages, Soliman, 2008a]

Our abstraction versus other abstractions

Reaction-thinking

Reaction network



stochastic

differential

approximation

correct abstraction

discrete

correct abstraction

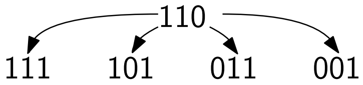
Boolean

Abstract simulation via FOBNN

correct abstraction

Conjecture: approximation

$A + B \rightarrow C$



110

111 101 011 001

[Fages, Soliman, 2008a]

Boolean network

Influence thinking

Outline

1. My method, its implementation and guarantees

SBML2BNET

STEP 1: Retrieve from the input reaction network

Structure: influence graph

Dynamics: Boolean transitions

- binarised time series from classic simulation of the ODEs
- abstract simulation of the ODEs

STEP 2: BN synthesis with `ASK&D-BN`

2. Evaluation of the approach
3. Link to other abstractions
4. Conclusion and perspectives

Conclusion and perspectives

Conclusion

Automatic synthesis of Boolean networks from a given reaction network, with guarentees. ✓

- Methodology
- Implementation (the SBML2BNET pipeline)
- Evaluation

Perspectives

1. *Ad hoc* solution to facilitate some analyses
   Make SBML2BNET easy to use
2. Better understanding of the formal relationship between reaction networks and Boolean network
   Two conjectures to investigate
3. Improve the Boolean networks synthesis methods when applied to wet data
   Investigate, in a controled environnement
   - overfitting to *the* sequence of configuration?
   - impact of the choice of the binarisation procedure and error measure

Publications

- J. Niehren et al. *Abstract Simulation of Reaction Networks via Boolean Networks* International Conference on Computational Methods in Systems Biology 2022
- A. Vaginay et al. *From Quantitative SBML Models to Boolean Networks* Complex Networks & Their Applications X 2022
- A. Vaginay et al. *From Quantitative SBML Models to Boolean Networks* Applied Network Science 2022
- A. Vaginay et al. *Automatic Synthesis of Boolean Networks from Biological Knowledge and Data* Optimization and Learning 2021
- A. Hirtz et al. *GPER Agonist G-1 Disrupts Tubulin Dynamics and Potentiates Temozolomide to Impair Glioblastoma Cell Proliferation* Cells 2021

Thank you for your attention.

References I

- [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

- [Hoops et al., 2006]
  S. Hoops et al.
  COPASI—a COmplex PAthway SImulator,
  Bioinformatics, vol. 22, pp. 3067–3074 2006

References II

- [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,

- [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

References III

- [Niehren et al., 2022]
  J. Niehren et al.
  Abstract Simulation of Reaction Networks via Boolean Networks
  CMSB2022: 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

- [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

References IV

- [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