From Chemical Reaction Networks to Boolean Networks, Automatically

Athénaïs Vaginay, Taha Boukhobza, Malika Smaïl-Tabbone

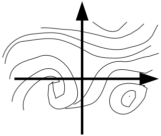
GT-BIOSS — December 3, 2021

Formalisms to model biological processes

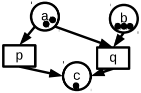
Chemical Reaction Network

```
e1 for A+B → C
e2 for C → D
e3 for _ → C
```

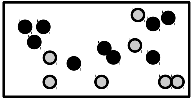
Differential equation



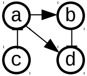
Petri Net



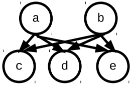
Agent-based model



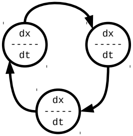
Boolean network



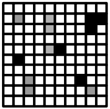
Bayesian network



Hybrid system



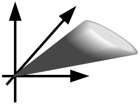
Cellular automata



Process algebra

```
((b(x,de)[E]) || (B(y, dI)[I]))
bh(x, dE) bh(y, dI) (E || I)
```

Constraint based model



Interacting state machine

Compartment based

Rule-based

...

Chemical Reaction Network (CRN)

$$\mathscr{M} = \{\mathscr{R}_i = e_i : R_i \xRightarrow{M_i} P_i\}_{i=1\ldots m}$$

Chemical Reaction Network (CRN)

$$\mathscr{M} = \{\mathscr{R}_i = e_i : R_i \xRightarrow{M_i} P_i\}_{i=1\ldots m}$$

$$\begin{aligned}
\mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\
\mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\
\mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}
\end{aligned}$$

Chemical Reaction Network (CRN)

$$\mathscr{M} = \{\mathscr{R}_i = e_i : R_i \overset{M_i}{\Longrightarrow} P_i\}_{i=1\ldots m}$$

$$\begin{aligned}
\mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\
\mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\
\mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}
\end{aligned}$$

reactants, products, modifiers, kinetics

Chemical Reaction Network (CRN)

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

$$\begin{aligned} \mathscr{R}_{\text{on}} &= \mathrm{k}_{\text{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\ \mathscr{R}_{\text{off}} &= \mathrm{k}_{\text{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\ \mathscr{R}_{\text{cat}} &= \mathrm{k}_{\text{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P} \end{aligned}$$

reactants, products, modifiers, kinetics

Chemical Reaction Network (CRN)

$$\mathscr{M} = \{\mathscr{R}_i = e_i : R_i \xRightarrow{M_i} P_i\}_{i=1\ldots m}$$

$$\begin{aligned}
\mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\
\mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\
\mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}
\end{aligned}$$

reactants, products, modifiers, kinetics

Chemical Reaction Network (CRN)

$$\mathscr{M} = \{\mathscr{R}_i = e_i : R_i \overset{M_i}{\Longrightarrow} P_i\}_{i=1\ldots m}$$

$$\begin{aligned}
\mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\
\mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\
\mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}
\end{aligned}$$

reactants, products, modifiers, kinetics

Chemical Reaction Network (CRN)

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

$$\begin{aligned}\mathscr{R}_{\mathrm{on}} &= \mathrm{k_{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}\\ \mathscr{R}_{\mathrm{off}} &= \mathrm{k_{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}\\ \mathscr{R}_{\mathrm{cat}} &= \mathrm{k_{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}\end{aligned}$$

reactants, products, modifiers, kinetics

Chemical Reaction Network (CRN)

$$\mathscr{M} = \{\mathscr{R}_i = e_i : R_i \overset{M_i}{\Longrightarrow} P_i\}_{i=1\ldots m}$$

$$\begin{aligned}
\mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\
\mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\
\mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}
\end{aligned}$$

reactants, products, modifiers, kinetics

CRN are versatile and well-studied. However...

Boolean Network (BN)

a set V of n components
Boolean status $\mathbb{B} = \{0, 1\}$
local update function $f_i : \mathbb{B}^n \rightarrow \mathbb{B}\ \forall i \in V$
$\neg$: "not"; $\vee$: "or"; $\wedge$: "and"

Boolean Network (BN)

a set V of n components
Boolean status $\mathbb{B} = \{0, 1\}$
local update function $f_i : \mathbb{B}^n \to \mathbb{B}\ \forall i \in V$
$\neg$: “not”; $\vee$: “or”; $\wedge$: “and”

Boolean network

$$\mathscr{B} = \begin{cases} f_{\mathsf{E}} := \neg \mathsf{E} \\ f_{\mathsf{ES}} := \mathsf{E} \wedge \mathsf{S} \\ f_{\mathsf{P}} := \mathsf{ES} \wedge (\neg \mathsf{E} \vee \mathsf{P}) \\ f_{\mathsf{S}} := \neg \mathsf{E} \end{cases}$$

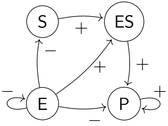
Boolean Network (BN)

a set V of n components
Boolean status $\mathbb{B} = \{0, 1\}$
local update function $f_i : \mathbb{B}^n \to \mathbb{B}\ \forall i \in V$
$\neg$: "not"; $\vee$: "or"; $\wedge$: "and"

Boolean network, its **interaction graph**

= structure

$$\mathscr{B} = \begin{cases} f_{\mathsf{E}} := \neg \mathsf{E} \\ f_{\mathsf{ES}} := \mathsf{E} \wedge \mathsf{S} \\ f_{\mathsf{P}} := \mathsf{ES} \wedge (\neg \mathsf{E} \vee \mathsf{P}) \\ f_{\mathsf{S}} := \neg \mathsf{E} \end{cases}$$

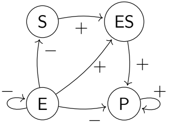


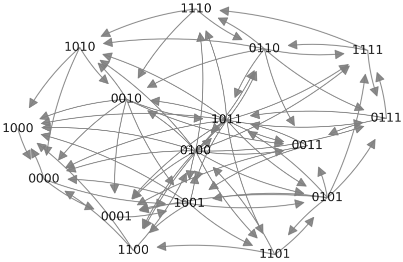
Boolean Network (BN)

a set V of n components
Boolean status $\mathbb{B} = \{0, 1\}$
local update function $f_i : \mathbb{B}^n \to \mathbb{B} \; \forall i \in V$
$\neg$: “not”; $\vee$: “or”; $\wedge$: “and”

Boolean network, its **interaction graph** and gen. asyn. **state transition graph**
= structure = dynamics

$$\mathscr{B} = \begin{cases} f_{\mathsf{E}} := \neg \mathsf{E} \\ f_{\mathsf{ES}} := \mathsf{E} \wedge \mathsf{S} \\ f_{\mathsf{P}} := \mathsf{ES} \wedge (\neg \mathsf{E} \vee \mathsf{P}) \\ f_{\mathsf{S}} := \neg \mathsf{E} \end{cases}$$





Goal

Synthesise Boolean networks starting from an existing chemical reaction network

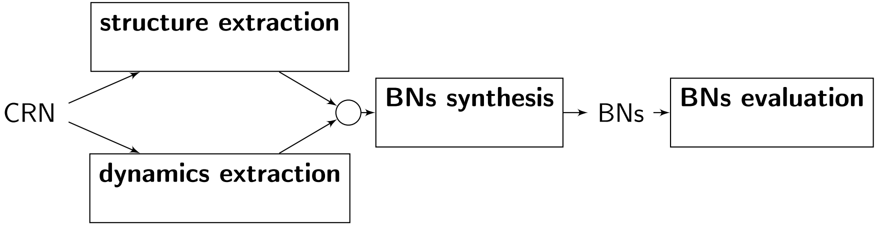
input: a chemical reaction network
output: a *set* of *compatible* Boolean networks

CRN → BNs

Goal, Proposed Pipeline

Synthesise Boolean networks starting from an existing chemical reaction network

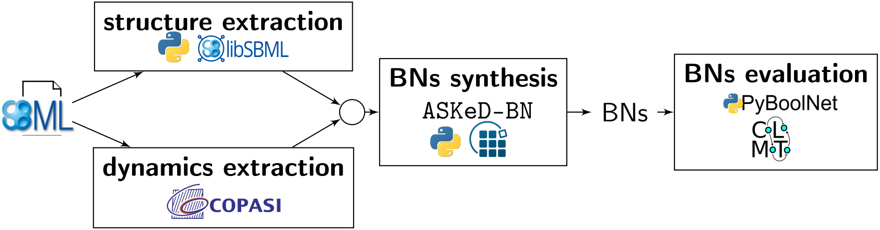
input: a chemical reaction network
output: a *set* of *compatible* Boolean networks



Goal, Proposed Pipeline, Implementation (=SBML2BN[1])

Synthesise Boolean networks starting from an existing chemical reaction network

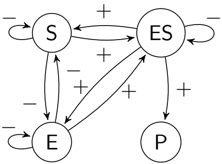
input: a chemical reaction network encoded in SBML
output: a *set* of *compatible* Boolean networks



[1] [Vaginay et al. CNA 2021]

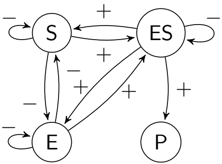
Details about the steps — structure extraction

$$\begin{aligned} \mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\ \mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\ \mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P} \end{aligned}$$



Details about the steps — structure extraction

$$\mathscr{R}_{\mathrm{on}} = \mathrm{k_{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}$$
$$\mathscr{R}_{\mathrm{off}} = \mathrm{k_{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}$$
$$\mathscr{R}_{\mathrm{cat}} = \mathrm{k_{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}$$



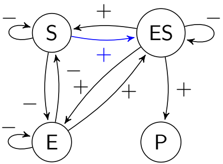
Syntactical Influence Graph (SIG)

1. **If** X is a reactant or an activator and Y disappears **then** X $\xrightarrow{-}$ Y
2. **If** X is an inhibitor and Y appears **then** X $\xrightarrow{-}$ Y
3. **If** X is a reactant or an activator and Y appears **then** X $\xrightarrow{+}$ Y
4. **If** X is an inhibitor and Y disappears **then** X $\xrightarrow{+}$ Y

[Fages et al. 2008]

Details about the steps — structure extraction

$$\begin{aligned}\mathscr{R}_{\text{on}} &= \mathrm{k}_{\text{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}\\ \mathscr{R}_{\text{off}} &= \mathrm{k}_{\text{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}\\ \mathscr{R}_{\text{cat}} &= \mathrm{k}_{\text{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2\cdot\mathsf{P}\end{aligned}$$



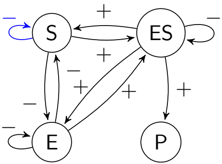
Syntactical Influence Graph (SIG)

1. **If** X is a reactant or an activator and Y disappears **then** X $\xrightarrow{-}$ Y
2. **If** X is an inhibitor and Y appears **then** X $\xrightarrow{-}$ Y
3. **If** X is a reactant or an activator and Y appears **then** X $\xrightarrow{+}$ Y
4. **If** X is an inhibitor and Y disappears **then** X $\xrightarrow{+}$ Y

[Fages et al. 2008]

Details about the steps — structure extraction

$$\begin{aligned} \mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\ \mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\ \mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P} \end{aligned}$$



Syntactical Influence Graph (SIG)

1. **If** X is a reactant or an activator and Y disappears **then** X $\xrightarrow{-}$ Y
2. **If** X is an inhibitor and Y appears **then** X $\xrightarrow{-}$ Y
3. **If** X is a reactant or an activator and Y appears **then** X $\xrightarrow{+}$ Y
4. **If** X is an inhibitor and Y disappears **then** X $\xrightarrow{+}$ Y

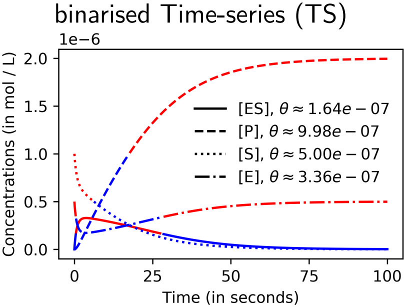
[Fages et al. 2008]

Details about the steps — dynamics extraction

$$\mathscr{R}_{\text{on}} = \text{k}_{\text{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}$$

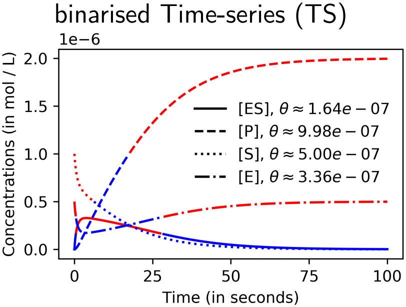
$$\mathscr{R}_{\text{off}} = \text{k}_{\text{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}$$

$$\mathscr{R}_{\text{cat}} = \text{k}_{\text{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}$$



Details about the steps — dynamics extraction

$$\mathscr{R}_{\text{on}} = \mathrm{k}_{\text{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}$$
$$\mathscr{R}_{\text{off}} = \mathrm{k}_{\text{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}$$
$$\mathscr{R}_{\text{cat}} = \mathrm{k}_{\text{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}$$

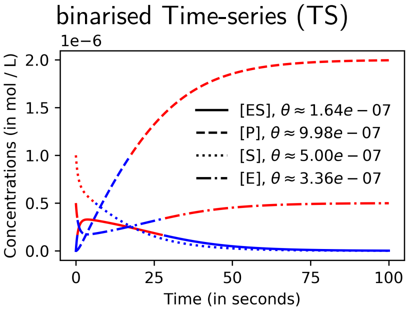


(1) reconstruct ODE

$$\forall i \in V : \frac{\mathrm{d}i}{\mathrm{d}t} = \sum_{r=1}^{m} f_r \times \delta_r(i)$$

Details about the steps — dynamics extraction

$$\begin{aligned}\mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}\\ \mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}\\ \mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}\end{aligned}$$



(1) reconstruct ODE

$$\forall i \in V : \frac{\mathrm{d}i}{\mathrm{d}t} = \sum_{r=1}^{m} f_r \times \delta_r(i)$$

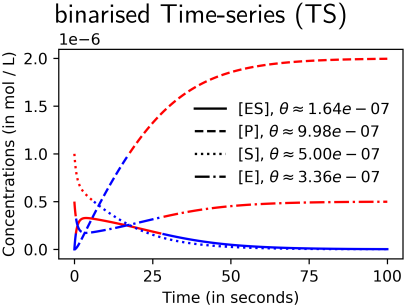
$$\frac{\mathrm{d}[\mathsf{ES}]}{\mathrm{d}t} = \underbrace{\mathrm{k}_{\mathrm{on}}[\mathsf{E}][\mathsf{S}] \times 1}_{\mathscr{R}_{\mathsf{on}}} + \underbrace{\mathrm{k}_{\mathrm{off}}[\mathsf{ES}] \times -1}_{\mathscr{R}_{\mathsf{off}}} + \underbrace{\mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] \times -1}_{\mathscr{R}_{\mathsf{cat}}}$$

Details about the steps — dynamics extraction

$$\mathscr{R}_{\mathrm{on}} = \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}$$

$$\mathscr{R}_{\mathrm{off}} = \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}$$

$$\mathscr{R}_{\mathrm{cat}} = \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}$$



(1) reconstruct ODE, (2) numerical simulation (parametrisation from the SBML model, duration of the simulation chosen by the user)

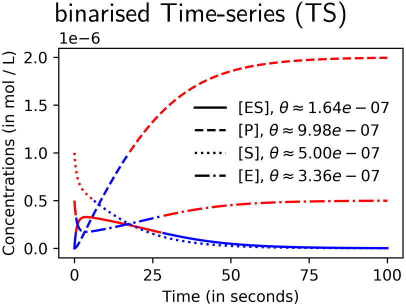
$$\forall i \in V : \frac{\mathrm{d}i}{\mathrm{d}t} = \sum_{r=1}^{m} f_r \times \delta_r(i)$$

Details about the steps — dynamics extraction

$$\mathscr{R}_{\mathrm{on}} = \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES}$$

$$\mathscr{R}_{\mathrm{off}} = \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E}$$

$$\mathscr{R}_{\mathrm{cat}} = \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}$$



(1) reconstruct ODE, (2) numerical simulation (parametrisation from the SBML model, duration of the simulation chosen by the user), (3) binarisation

$$\forall i \in V : \frac{\mathrm{d}i}{\mathrm{d}t} = \sum_{r=1}^{m} f_r \times \delta_r(i)$$

Details about the steps — BNs Synthesis

Find $f_i : \mathbb{B}^n \rightarrow \mathbb{B}\ \forall i \in V$ *compatible* with structural constraints (= domain) and dynamical constraints (= to reproduce)

Details about the steps — BNs Synthesis

Find $f_i : \mathbb{B}^n \to \mathbb{B}\ \forall i \in V$ *compatible* with structural constraints (= domain) and dynamical constraints (= to reproduce)

Disjunctive Normal Form (DNF) = disjunction of conjunctions

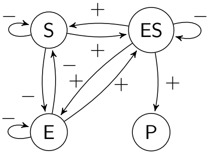
$$f_{\mathsf{ES}} := \underbrace{(\neg\mathsf{ES} \land \mathsf{E} \land \mathsf{S})}_{\text{conj. 1}} \lor \underbrace{(\mathsf{E} \land \neg\mathsf{P})}_{\text{conj. 2}}$$

Details about the steps — BNs Synthesis

Find $f_i : \mathbb{B}^n \to \mathbb{B}\ \forall i \in V$ *compatible* with structural constraints (= domain) and dynamical constraints (= to reproduce)

Disjunctive Normal Form (DNF) = disjunction of conjunctions

$$f_{\mathsf{ES}} := \underbrace{(\neg \mathsf{ES} \land \mathsf{E} \land \mathsf{S})}_{\text{conj. 1}} \lor \underbrace{(\mathsf{E} \land \neg \mathsf{P})}_{\text{conj. 2}}$$



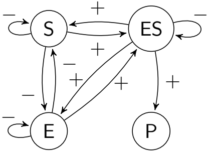
Hard constraint: j appears positively (resp. negatively) in f_i iff $j \xrightarrow{+} i$ (resp. $j \xrightarrow{-} i$)

Details about the steps — BNs Synthesis

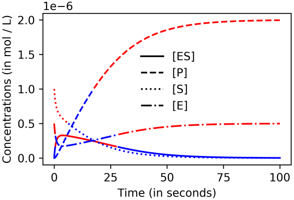
Find $f_i : \mathbb{B}^n \to \mathbb{B}\ \forall i \in V$ *compatible* with structural constraints (= domain) and dynamical constraints (= to reproduce)

Disjunctive Normal Form (DNF) = disjunction of conjunctions

$$f_{\mathsf{ES}} := \underbrace{(\neg \mathsf{ES} \wedge \mathsf{E} \wedge \mathsf{S})}_{\text{conj. 1}} \vee \underbrace{(\mathsf{E} \wedge \neg \mathsf{P})}_{\text{conj. 2}}$$



Hard constraint: j appears positively (resp. negatively) in f_i iff $j \xrightarrow{+} i$ (resp. $j \xrightarrow{-} i$)

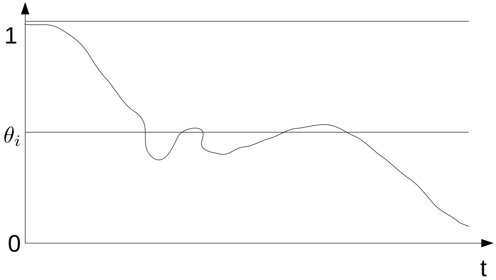


Soft constraint: f_i minimise the unexplained transitions → penalty if:

$c \in f_i$ observed at t but $i_{t+1} = 0$

$i_{t+1} = 1$ but no $c \in f_i$ observed at t

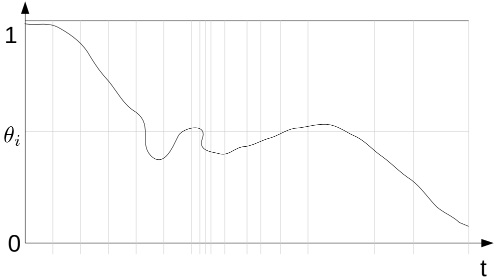
Details about the steps — BNs Synthesis: soft constraint



i_t: observation if i at time t

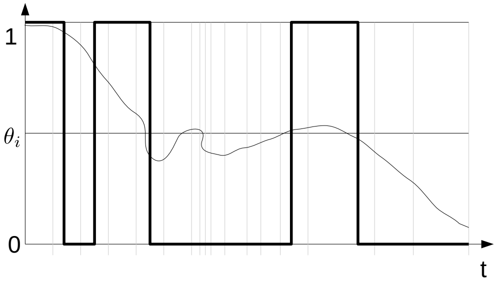
θ_i: binarisation threshold for i

Details about the steps — BNs Synthesis: soft constraint



i_t: observation if i at time t
θ_i: binarisation threshold for i
T: # time steps

Details about the steps — BNs Synthesis: soft constraint

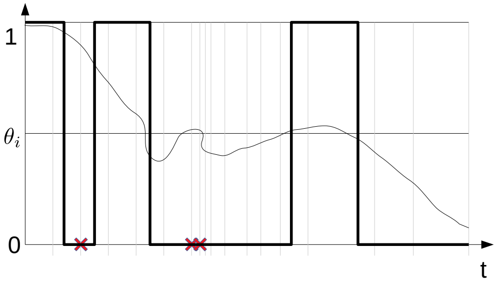


i_t: observation if i at time t

θ_i: binarisation threshold for i

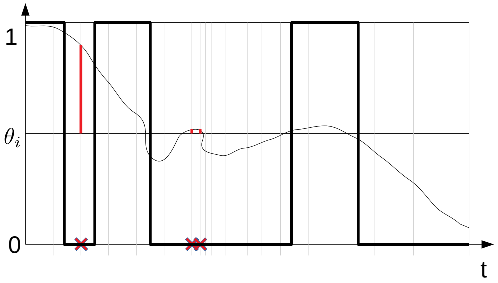
T: # time steps

Details about the steps — BNs Synthesis: soft constraint



i_t: observation if i at time t
θ_i: binarisation threshold for i
T: # time steps
$\mathscr{U}$: set of unexplained time steps

Details about the steps — BNs Synthesis: soft constraint



i_t: observation if i at time t
θ_i: binarisation threshold for i
T: # time steps
$\mathscr{U}$: set of unexplained time steps

minimise the Mean Absolute Error (ideally 0)

$$\mathrm{MAE}_{f_i} = \frac{\sum_{t \in \mathscr{U}_{f_i}} |\theta_i - i_t|}{T}$$

Details about the steps — BNs Synthesis

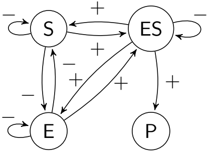
Logic program (Answer-Set Programming) `ASKeD-BN`[2]
constraints: structural (= domain) and dynamical (= to reproduce)

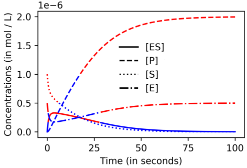
output: exhaustive set of BNs whose interaction graph is a subgraph of the SIG and their dynamics minimise the unexplained transitions.

[2] [Vaginay et al. OLA 2021]

Details about the steps — BNs Synthesis

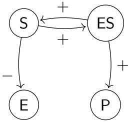
Logic program (Answer-Set Programming) `ASKeD-BN`[2]
constraints: structural (= domain) and dynamical (= to reproduce)

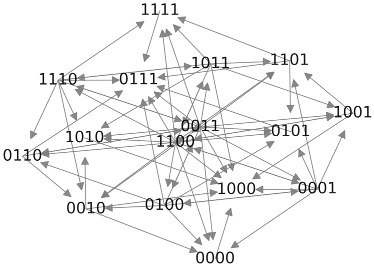




output: exhaustive set of BNs whose interaction graph is a subgraph of the SIG and their dynamics minimise the unexplained transitions.

$$\mathscr{B}_1 = \begin{cases} f_{\mathrm{ES}} := \mathrm{S} \\ f_{\mathrm{P}} := \mathrm{ES} \\ f_{\mathrm{S}} := \mathrm{ES} \\ f_{\mathrm{E}} := \neg \mathrm{S} \end{cases}$$





[2] [Vaginay et al. OLA 2021]

Details about the steps — BNs evaluation

- hard constraint: assert the interaction graph of synthesised BNs are subgraph of the SIG ✓(by construction)
- compute the "coverage proportion" of each synthesised BN

$$= \frac{\#\ \text{recovered transitions}}{\#\ \text{transitions observed}}$$

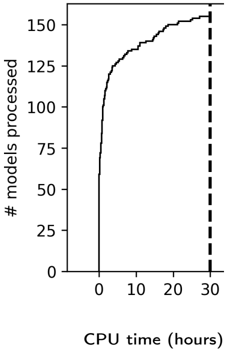
(= 1 if the BN reproduces perfectly the sequence of configurations)

Pipeline evaluation and results

Pipeline ran on 209 SBML models from Biomodels.
# components: 2 – 60
# parents: 10 max
→ simple to medium complexity

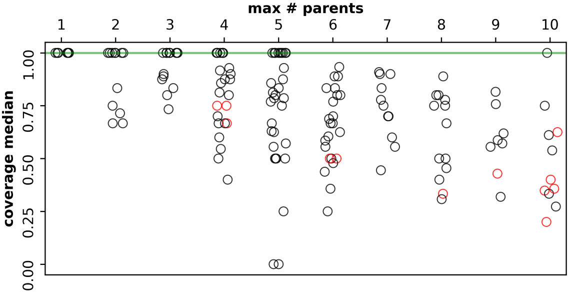
Results about the runtime and the coverage of the BNs synthesised for each input SBML model.

Results — runtime

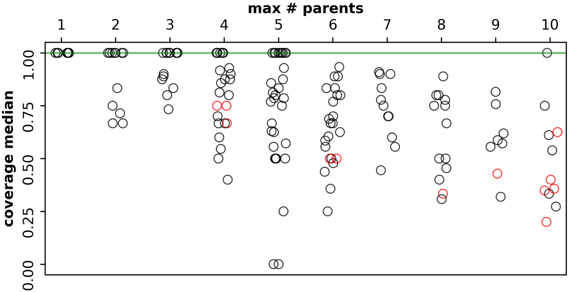


$\sim$75% (=155/209) $\leq$ 30 hours
$\sim$ half $\leq$ 30 mins

Results — coverage proportion

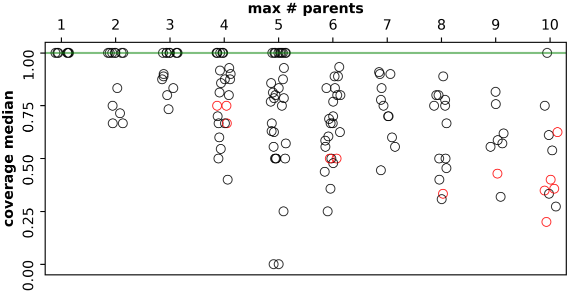


Results — coverage proportion



Med=0.77; Var=0, even when > 1 BN synthesised (except for 12 models)

Results — coverage proportion



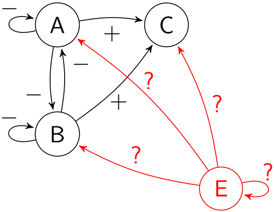
Med=0.77; Var=0, even when > 1 BN synthesised (except for 12 models)

Loss of performance when max # parents increases.

Impact of SBML inconsistencies on structure extraction

Ex. BIOMD n°44: 1 BN generated; coverage=0.55
some kinetics use components not listed in the reactants nor modifiers → incomplete SIG (missing parents)

$f(\mathsf{A}, \mathsf{B}, \mathsf{E}) : \mathsf{A} + \mathsf{B} \Rightarrow \mathsf{C}$

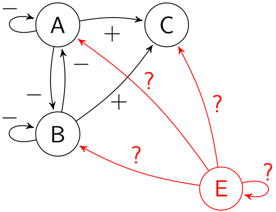


[3] [Fages et al. 2012]

Impact of SBML inconsistencies on structure extraction

Ex. BIOMD n°44: 1 BN generated; coverage=0.55
some kinetics use components not listed in the reactants nor modifiers → incomplete SIG (missing parents)

$$f(\mathsf{A},\mathsf{B},\mathsf{E}) : \mathsf{A} + \mathsf{B} \Rightarrow \mathsf{C}$$



> 60% of SBML models from Biomodels are not "well-formed"[3], but some can be fixed → add a step in the pipeline

[3] [Fages et al. 2012]

Conclusion

Automatic transformation of a CRN into a set of BNs
SBML2BN = a proof of concept with possible improvements...

- for even more complex models (models with > 10 parents)
- fix not well-formed SBML models
- take more constraints into account (fixed-points)
- *correct* abstraction?

Thanks for your attention.

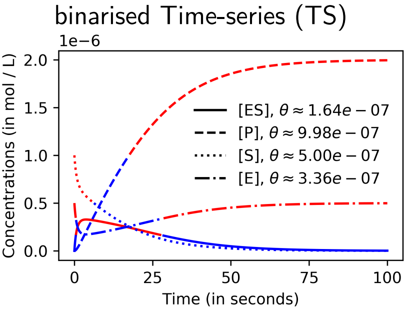
Hope to $\{$see, read$\}$ you.

`athenais.vaginay@loria.fr`

Thanks for the opportunity to present our work! :)

Details about the steps — dynamics extraction

$$\begin{aligned}
\mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\
\mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\
\mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P}
\end{aligned}$$



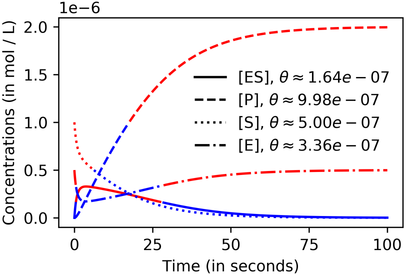
numerical simulation of the reconstructed ODE (parametrisation from the SBML model, duration of the simulation chosen by the user) + binarisation

$$\begin{aligned}
\frac{\mathrm{d}[\mathrm{E}]}{\mathrm{d}t} &= -\mathrm{k}_{\mathrm{on}}[\mathrm{E}][\mathrm{S}] + \mathrm{k}_{\mathrm{off}}[\mathrm{ES}] + \mathrm{k}_{\mathrm{cat}}[\mathrm{ES}] \\
\frac{\mathrm{d}[\mathrm{ES}]}{\mathrm{d}t} &= \mathrm{k}_{\mathrm{on}}[\mathrm{E}][\mathrm{S}] - \mathrm{k}_{\mathrm{off}}[\mathrm{ES}] - \mathrm{k}_{\mathrm{cat}}[\mathrm{ES}] \\
\frac{\mathrm{d}[\mathrm{P}]}{\mathrm{d}t} &= 2\,\mathrm{k}_{\mathrm{cat}}[\mathrm{ES}] \\
\frac{\mathrm{d}[\mathrm{S}]}{\mathrm{d}t} &= -\mathrm{k}_{\mathrm{on}}[\mathrm{E}][\mathrm{S}] + \mathrm{k}_{\mathrm{off}}[\mathrm{ES}]
\end{aligned}$$

Details about the steps — dynamics extraction

$$\begin{aligned} \mathscr{R}_{\mathrm{on}} &= \mathrm{k}_{\mathrm{on}}[\mathsf{S}][\mathsf{E}] : \mathsf{S} + \mathsf{E} \Rightarrow \mathsf{ES} \\ \mathscr{R}_{\mathrm{off}} &= \mathrm{k}_{\mathrm{off}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{S} + \mathsf{E} \\ \mathscr{R}_{\mathrm{cat}} &= \mathrm{k}_{\mathrm{cat}}[\mathsf{ES}] : \mathsf{ES} \Rightarrow \mathsf{E} + 2 \cdot \mathsf{P} \end{aligned}$$

binarised Time-series (TS)



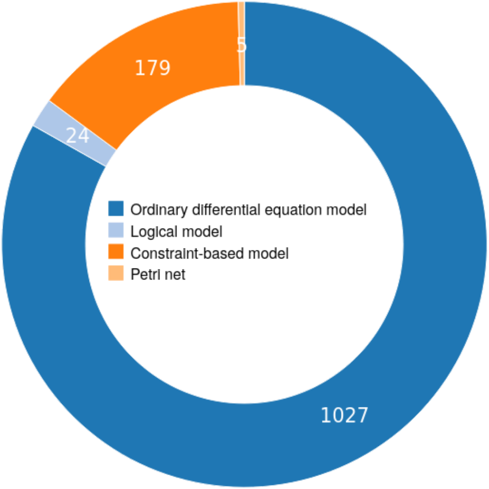
---

numerical simulation of the reconstructed ODE (parametrisation from the SBML model, duration of the simulation chosen by the user) + binarisation

$$\begin{aligned} \frac{\mathrm{d[E]}}{\mathrm{d}t} &= -\mathrm{k}_{\mathrm{on}}[\mathrm{E}][\mathrm{S}] + \mathrm{k}_{\mathrm{off}}[\mathrm{ES}] + \mathrm{k}_{\mathrm{cat}}[\mathrm{ES}] \\ \frac{\mathrm{d[ES]}}{\mathrm{d}t} &= \mathrm{k}_{\mathrm{on}}[\mathrm{E}][\mathrm{S}] - \mathrm{k}_{\mathrm{off}}[\mathrm{ES}] - \mathrm{k}_{\mathrm{cat}}[\mathrm{ES}] \\ \frac{\mathrm{d[P]}}{\mathrm{d}t} &= 2\,\mathrm{k}_{\mathrm{cat}}[\mathrm{ES}] \\ \frac{\mathrm{d[S]}}{\mathrm{d}t} &= -\mathrm{k}_{\mathrm{on}}[\mathrm{E}][\mathrm{S}] + \mathrm{k}_{\mathrm{off}}[\mathrm{ES}] \end{aligned}$$

CRNs in Biomodels

1039 manually curated models (data from 08th Nov. 2021) available in the System Biology Markup Language (SBML)



Semantics of CRN $\mathscr{M} = \{f(\mathsf{A}, \mathsf{B}) : \mathsf{A} + \mathsf{B} \Rightarrow \mathsf{C}\}$

Stochastic

Continuous Time Markov Chain

numbers of molecules

$\mathsf{A}, \mathsf{B} \xrightarrow{p(S_i), t(S_i)} \mathsf{C}{++}, \mathsf{A}{--}, \mathsf{B}{--}$

Continuous

Ordinary Differential Equations

concentrations over time

$\forall i \in V : \frac{\mathrm{d}i}{\mathrm{d}t} = \sum_{r=1}^{m} f_r \delta_r(i)$

Discrete

Petri net

numbers of molecules

$\mathsf{A}, \mathsf{B} \rightarrow \mathsf{C}{++}, \mathsf{A}{--}, \mathsf{B}{--}$

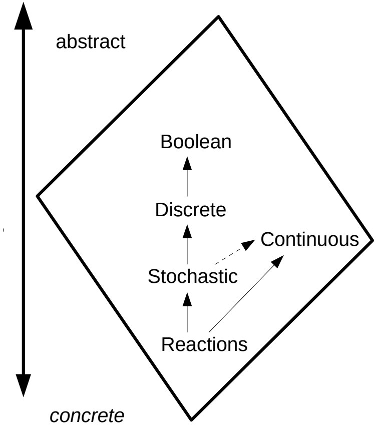
Boolean

Asynchronous Transition System

presence / absence of molecules

$\mathsf{A} \wedge \mathsf{B} \rightarrow \mathsf{C} \wedge \mathsf{A}/\neg\mathsf{A} \wedge \mathsf{B}/\neg\mathsf{B}$

Hierachy of CRN semantics



Galois connection between the syntactical, stochastic, discrete and Boolean semantics[4]

If a behaviour is not possible in the Boolean semantics, it is not possible in the stochastic semantics for any reaction rates

Under large number conditions, the ODE semantics approximates the mean stochastic behaviour[5].

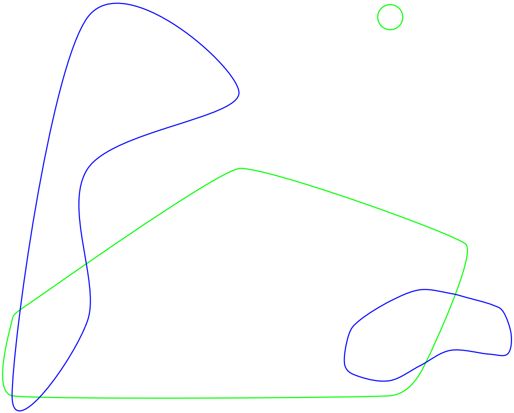
Slide adapted from Francois Fages' presentation Bioregul 2019

[4] [Fages & Soliman 2006, 2008]

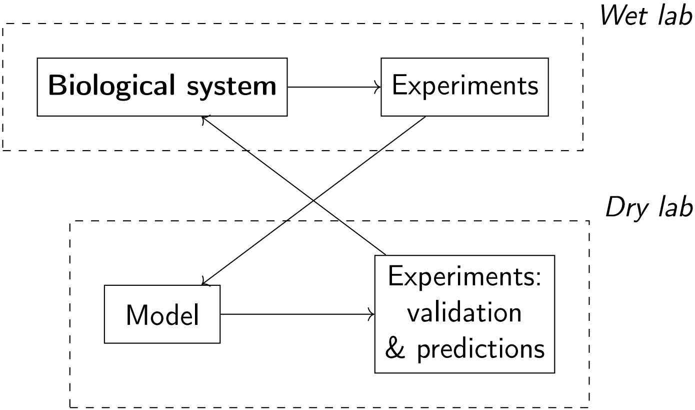
[5] [Gillespie 71]

Details about the steps — BNs Synthetis

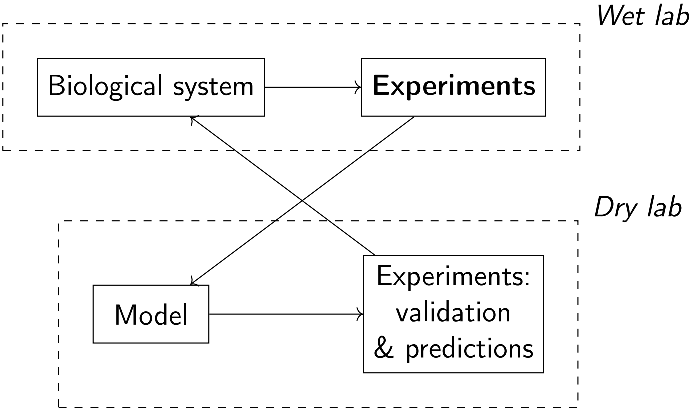
Method: `ASKeD-BN` [Vaginay et al. OLA 2021] (Answer-Set Programming)



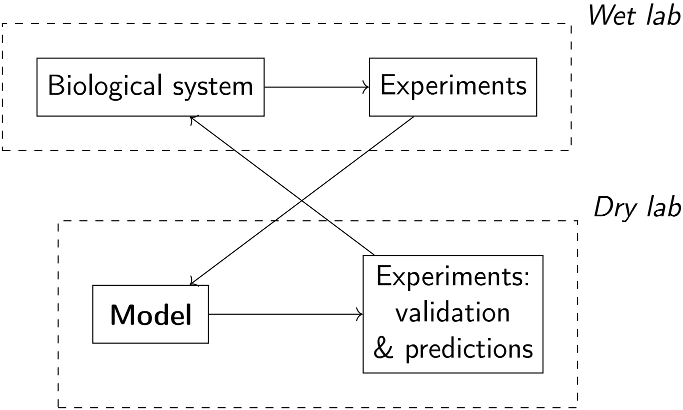
Modelling of biological systems



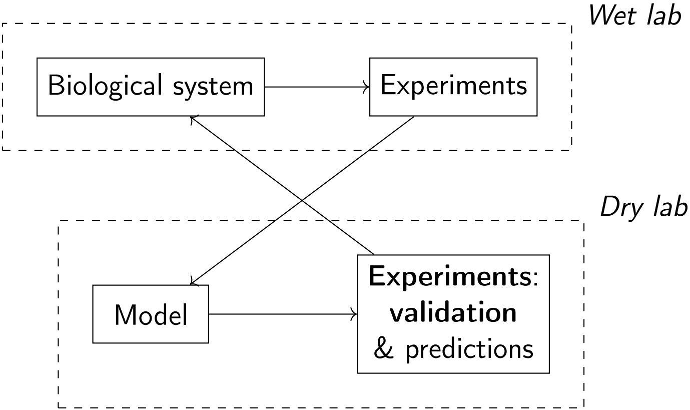
Modelling of biological systems



Modelling of biological systems



Modelling of biological systems



Wet lab

Biological system

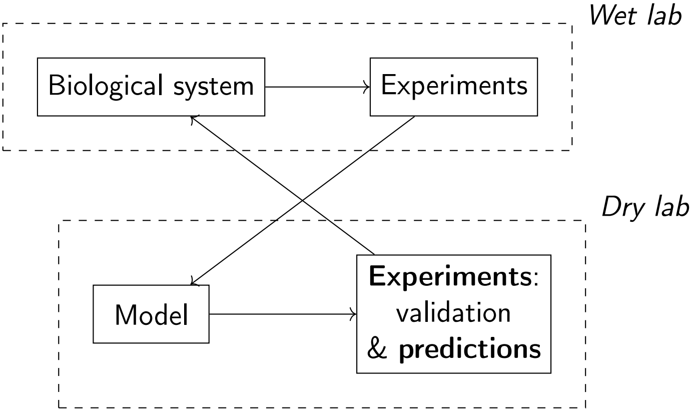
Experiments

Dry lab

Model

Experiments: **validation** & predictions

Modelling of biological systems



Wet lab

Biological system

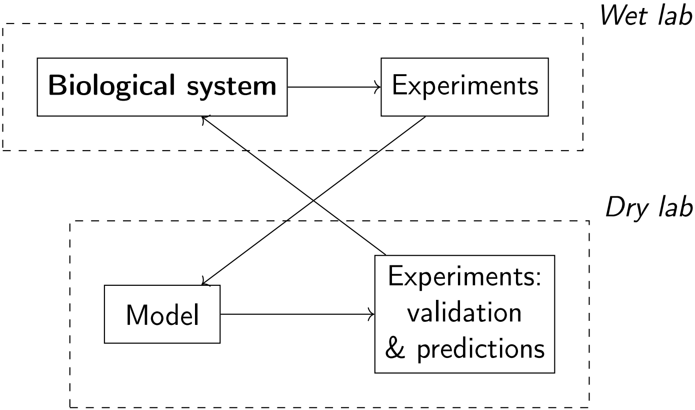
Experiments

Dry lab

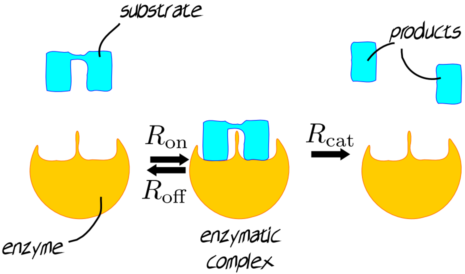
Model

Experiments: validation & **predictions**

Modelling of biological systems

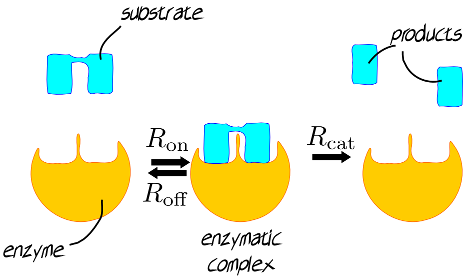


Modelling of biological systems — Example: enzymatic reaction



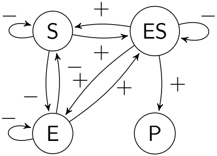
$$\mathsf{S} + \mathsf{E} \rightleftharpoons \mathsf{ES} \longrightarrow \mathsf{E} + 2 \cdot \mathsf{P}$$

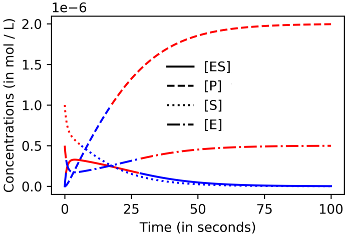
Modelling of biological systems — Example: enzymatic reaction



$$\mathsf{S} + \mathsf{E} \rightleftharpoons \mathsf{ES} \longrightarrow \mathsf{E} + 2 \cdot \mathsf{P}$$

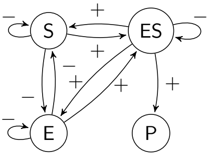
constraints: known structure (= domain) and dynamics (= to reproduce)

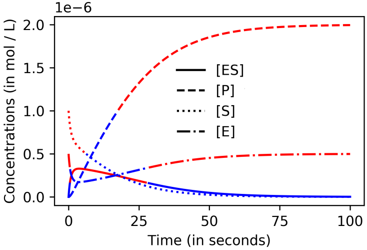




Boolean Networks Synthesis

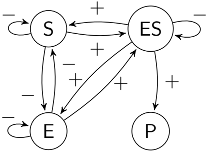
constraints: known structure (= domain) and dynamics (= to reproduce)

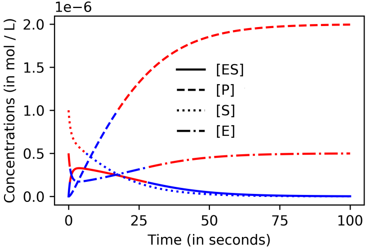




Boolean Networks Synthesis

constraints: known structure (= domain) and dynamics (= to reproduce)

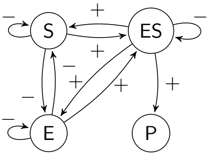


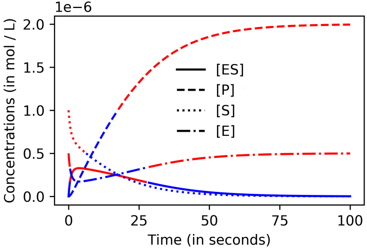


output: (exhaustive set of) BNs compatible with the constraints

Boolean Networks Synthesis

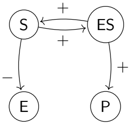
constraints: known structure (= domain) and dynamics (= to reproduce)

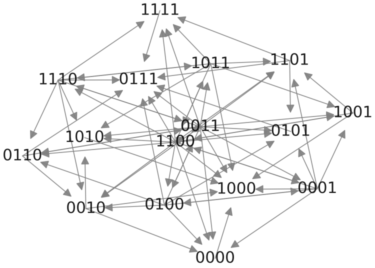




output: (exhaustive set of) BNs compatible with the constraints

$$\mathscr{B}_1 = \begin{cases} f_{\mathrm{ES}} := \mathrm{S} \\ f_{\mathrm{P}} := \mathrm{ES} \\ f_{\mathrm{S}} := \mathrm{ES} \\ f_{\mathrm{E}} := \neg \mathrm{S} \end{cases}$$





And then what?

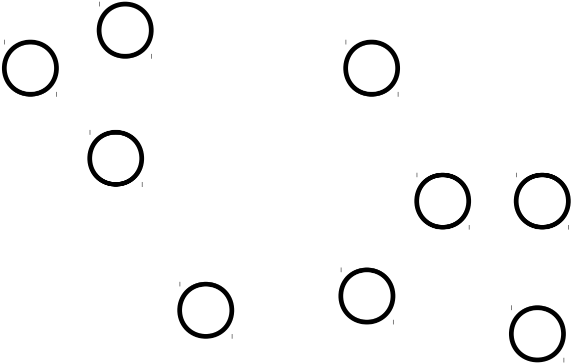
Analysis on Boolean networks (even large ones !)

- reachability, attractors
- control
- . . .

Despite their simplicity, BNs can fit complex biological phenomena.
Used for: cell cycle, cancer (breast, bladder, . . .), . . .

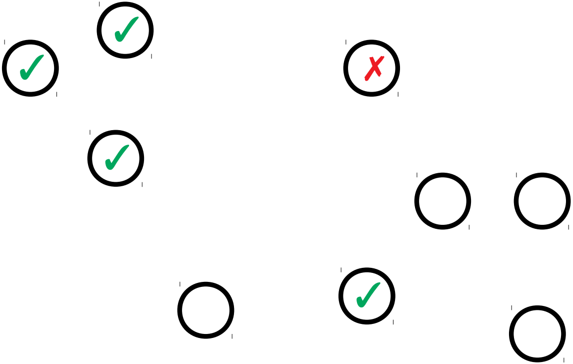
BN control

Possible configurations of a system



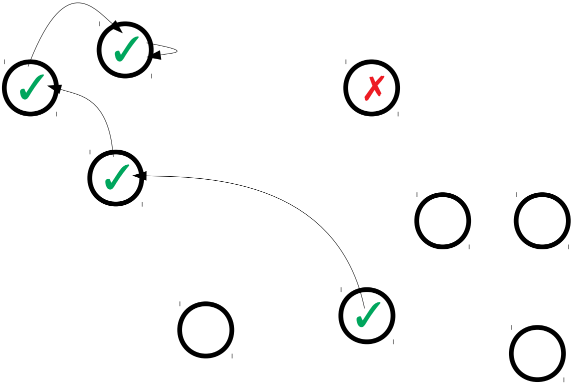
BN control

some are “normal”, some are “pathological”



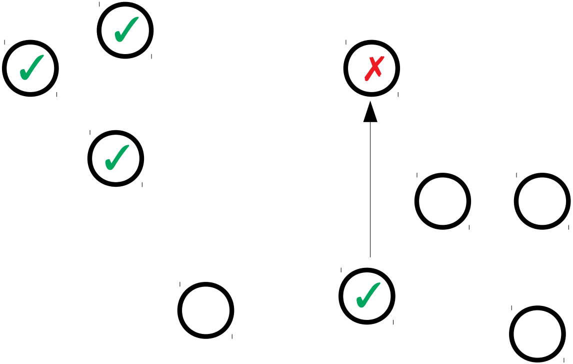
BN control

behaviour of a **sane** system goes through normal configurations...



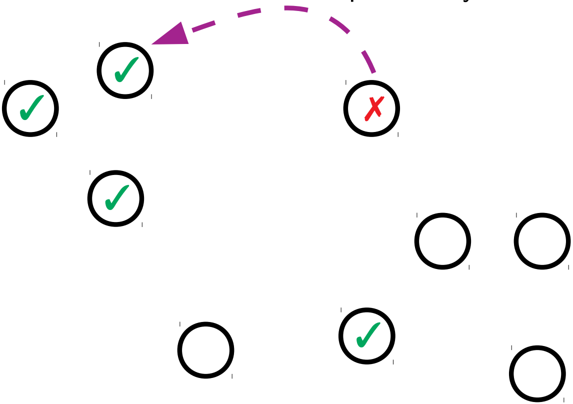
BN control

... while a **broken** system gets stuck in a pathological configuration



BN control

Which interventions can repair the system?

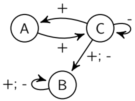


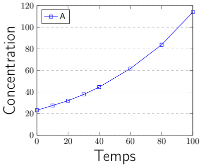
Conclusion

Boolean networks:

- simple, yet powerful formalism to study biological processes
- simple to set up: their construction requires very few data compared to other formalisms
- once built, one can run prediction analysis, control, etc. . .

Structural & dynamical constraints

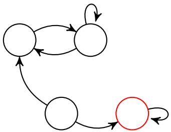




Boolean Networks(BN)

$$\begin{cases} f_{\mathsf{A}} := \mathsf{C} \\ f_{\mathsf{B}} := \mathsf{B} \oplus \mathsf{C} \\ f_{\mathsf{C}} := \mathsf{A} \land \neg \mathsf{C} \end{cases}$$

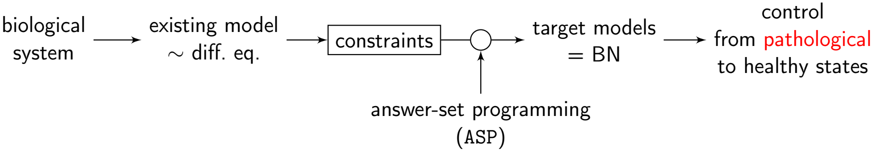
Control



PhD Project

1er octobre 2018 → juin 2022

“Selection and analysis of models for biology using knowledge on the domain; application to pathological systems.”



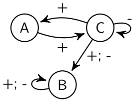
biological system → existing model ∼ diff. eq. → constraints → target models = BN → control from pathological to healthy states

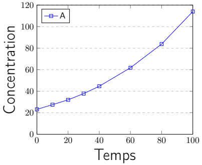
answer-set programming (ASP)

Structural & dynamical constraints

Boolean Networks (BN)

Control





$$\begin{cases} f_{\mathsf{A}} := \mathsf{C} \\ f_{\mathsf{B}} := \mathsf{B} \oplus \mathsf{C} \\ f_{\mathsf{C}} := \mathsf{A} \wedge \neg \mathsf{C} \end{cases}$$