From Quantitative SBML to Boolean Networks

COMPLEX
NETWORKS

Athénaïs Vaginay, Taha Boukhobza, Malika Smaïl-Tabbone
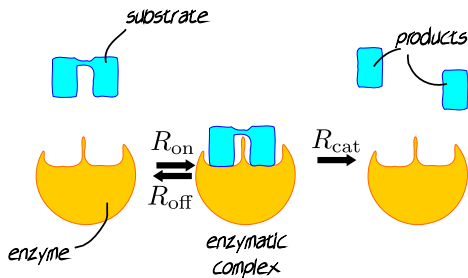
November 30 — December 2, 2021

Outline

- ▶ Modelling of biological systems
  - ▶ generalities
  - ▶ quantitative SBML models
  - ▶ Boolean Networks
- ▶ SBML2BN pipeline
- ▶ Empirical validation, results
- ▶ Discussion, Conclusion

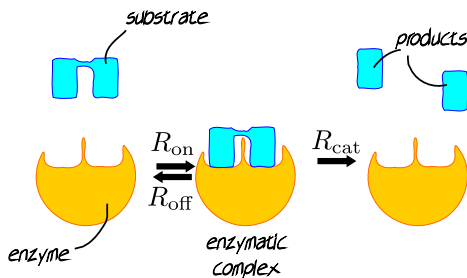# Modelling of biological systems

Modelling of biological systems

Example of a simple enzymatic reaction:

Modelling of biological systems
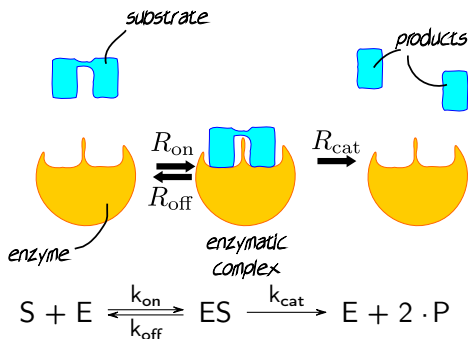
Example of a simple enzymatic reaction:

reactants of R_x species consumed by reaction R_x

products of R_x species produced by reaction R_x

Modelling of biological systems

Example of a simple enzymatic reaction:

reactants of R_x species consumed by reaction R_x

products of R_x species produced by reaction R_x

Modelling of biological systems

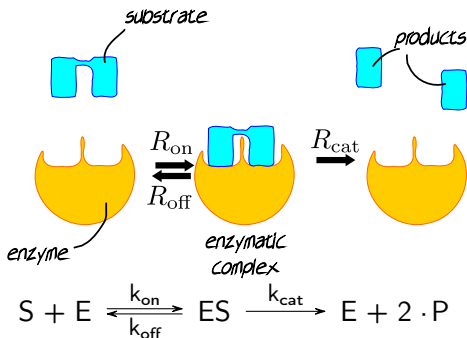Example of a simple enzymatic reaction:

reactants of R_x species consumed by reaction R_x

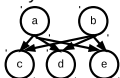products of R_x species produced by reaction R_x

speed constants proportionality coefficient between the amount of the reactants and the rate of the reaction

Existing formalisms...

Boolean network

Bayesian network

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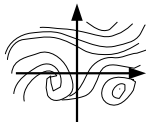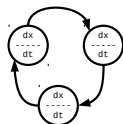```
((b(x,de)[E]) || (B(y, dI)[I]))
bh(x, dE) bh(y, dI) (E || I)
```
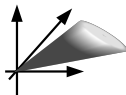
Differential equation

Constraint based model

Hybrid systems

Petri Nets
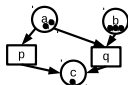
Agent-based model
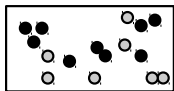
Cellular automata

Interacting state machine
Compartment based
Rule based

...

# Systems Biology Markup Language

chemical reactions network = set of reactions

$$R_{\mathrm{on}} : \mathsf{S + E} \xrightarrow{\mathsf{k_{on}}} \mathsf{ES}$$

$$R_{\mathrm{off}} : \mathsf{ES} \xrightarrow{\mathsf{k_{off}}} \mathsf{S + E}$$

$$R_{\mathrm{cat}} : \mathsf{ES} \xrightarrow{\mathsf{k_{cat}}} \mathsf{E + 2 \cdot P}$$
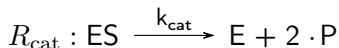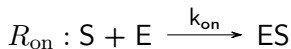
complete quantitative SBML models
$\rightarrow$ after ODE reconstruction, we can get concentration of
components over time

# However. . .

. . . in some cases, qualitative models such as Boolean networks would be more suited because their simplicity make them easy to study (attractor, control, . . . ).

# Boolean Networks (BNs)

BN = set of $n$ *transition functions* $\mathbb{B}^n \to \mathbb{B}$  $\mathbb{B} = \{0, 1\}$

logical operators = { $\neg$: "not"; $\vee$: "or"; $\wedge$: "and" }

# Boolean Networks (BNs)

$$\text{BN} = \text{set of } n \text{ transition functions } \mathbb{B}^n \to \mathbb{B} \qquad \mathbb{B} = \{0, 1\}$$
$$\text{logical operators} = \{ \neg: \text{"not"}; \vee: \text{"or"}; \wedge: \text{"and"} \}$$

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 Networks (BNs)

$$\text{BN} = \text{set of } n \text{ transition functions } \mathbb{B}^n \to \mathbb{B} \qquad \mathbb{B} = \{0, 1\}$$
$$\text{logical operators} = \{ \neg: \text{"not"}; \vee: \text{"or"}; \wedge: \text{"and"} \}$$
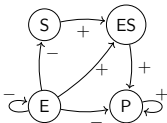
Boolean network, its interaction graph
$$= \text{structure}$$

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

# Boolean Networks (BNs)

BN = set of $n$ *transition functions* $\mathbb{B}^n \to \mathbb{B}$ $\qquad \mathbb{B} = \{0, 1\}$
logical operators = { ¬: "not"; ∨: "or"; ∧: "and" }

Boolean network, its interaction graph and gen. asyn. state transition graph
$\qquad$ = structure $\qquad\qquad$ = dynamics

$$\mathscr{B} = \begin{cases} f_\mathsf{E} := \neg\mathsf{E} \\ f_\mathsf{ES} := \mathsf{E} \wedge \mathsf{S} \\ f_\mathsf{P} := \mathsf{ES} \wedge (\neg\mathsf{E} \vee \mathsf{P}) \\ f_\mathsf{S} := \neg\mathsf{E} \end{cases}$$
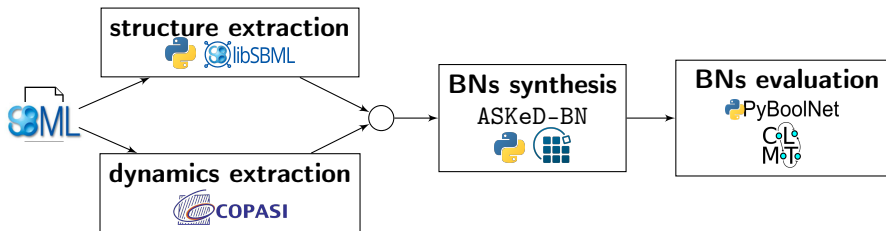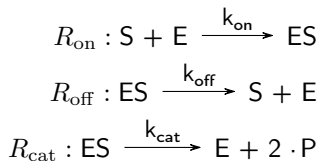
# The proposed pipeline: SBML2BN

Goal: Synthesise Boolean networks starting from a preexisting SBML model

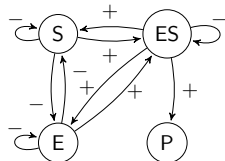**input**: a *complete* quantitative SBML model
**output**: a *set* of *compatible* Boolean networks
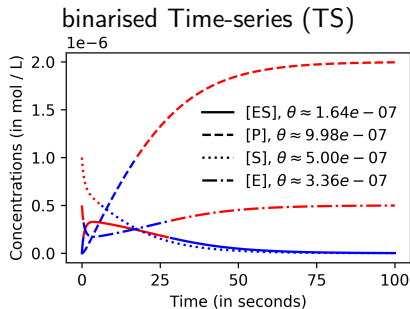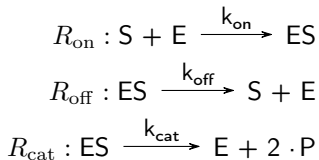
# Details about the steps — structure extraction

Prior Knowledge Network (PKN)

$$R_{\text{on}} : \mathsf{S} + \mathsf{E} \xrightarrow{\ \mathsf{k_{on}}\ } \mathsf{ES}$$

$$R_{\text{off}} : \mathsf{ES} \xrightarrow{\ \mathsf{k_{off}}\ } \mathsf{S} + \mathsf{E}$$

$$R_{\text{cat}} : \mathsf{ES} \xrightarrow{\ \mathsf{k_{cat}}\ } \mathsf{E} + 2 \cdot \mathsf{P}$$



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** $X \xrightarrow{+} Y$

# Details about the steps — dynamics extraction

$$R_{\mathrm{on}} : \mathsf{S} + \mathsf{E} \xrightarrow{\mathsf{k_{on}}} \mathsf{ES}$$

$$R_{\mathrm{off}} : \mathsf{ES} \xrightarrow{\mathsf{k_{off}}} \mathsf{S} + \mathsf{E}$$

$$R_{\mathrm{cat}} : \mathsf{ES} \xrightarrow{\mathsf{k_{cat}}} \mathsf{E} + 2 \cdot \mathsf{P}$$



binarised Time-series (TS)

[ES], $\theta \approx 1.64e - 07$
[P], $\theta \approx 9.98e - 07$
[S], $\theta \approx 5.00e - 07$
[E], $\theta \approx 3.36e - 07$

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

$$\frac{\mathrm{d[E]}}{\mathrm{d}t} = -\mathrm{k_{on}[E][S]} + \mathrm{k_{off}[ES]} + \mathrm{k_{cat}[ES]}$$
$$\frac{\mathrm{d[ES]}}{\mathrm{d}t} = \mathrm{k_{on}[E][S]} - \mathrm{k_{off}[ES]} - \mathrm{k_{cat}[ES]}$$
$$\frac{\mathrm{d[P]}}{\mathrm{d}t} = 2\,\mathrm{k_{cat}[ES]}$$
$$\frac{\mathrm{d[S]}}{\mathrm{d}t} = -\mathrm{k_{on}[E][S]} + \mathrm{k_{off}[ES]}$$

# Details about the steps — BNs Synthetis

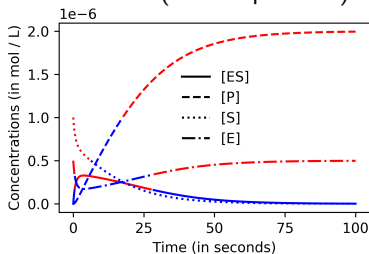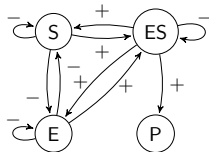Tool: `ASKeD-BN` [Vaginay et al. OLA 2021] (Answer-Set Programming)
**inputs**: a PKN (= domain) and a binarised TS (= to reproduce)
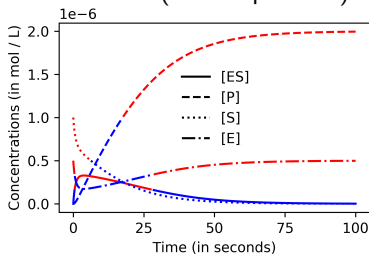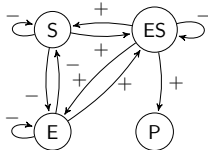
# Details about the steps — BNs Synthetis

Tool: `ASKeD-BN` [Vaginay et al. OLA 2021] (Answer-Set Programming)
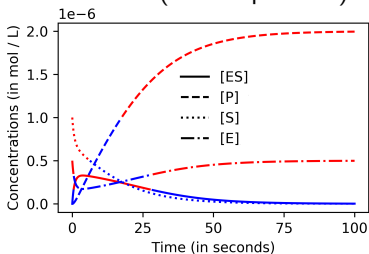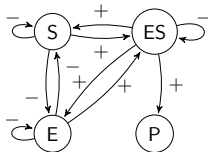**inputs**: a PKN (= domain) and a binarised TS (= to reproduce)



**output**: exhaustive set of BNs whose interaction graph is a subgraph of the PKN and their dynamics minimise the number of unexplained transitions.
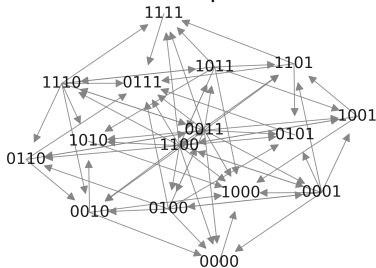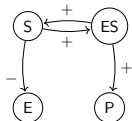
# Details about the steps — BNs Synthetis

Tool: `ASKeD-BN` [Vaginay et al. OLA 2021] (Answer-Set Programming)
**inputs**: a PKN (= domain) and a binarised TS (= to reproduce)



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

$$\mathscr{B}_1 = \begin{cases} f_{ES} := S \\ f_P := ES \\ f_S := ES \\ f_E := \neg S \end{cases}$$

# Details about the steps — BNs evaluation

▶ hard constraint: assert the interaction graph of synthesised BNs are subgraph of the PKN ✓(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

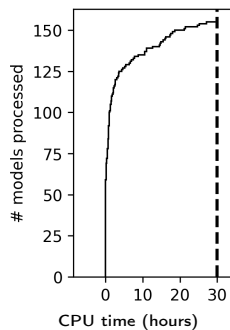Pipeline ran on 209 SBML models from Biomodels.
# components: 2 – 60
# **parents: 10 max**
$\rightarrow$ simple to medium complexity

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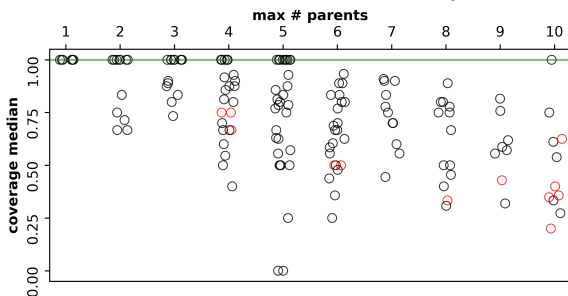for each input SBML model.

# Results — runtime



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

# Results — coverage proportion

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



Loss of performance when # max parents increases.

But also if models not "well-formed"? [Fages et al. 2012]
Ex. BIOMD n°44: 1 BN generated; coverage=0.55
But some kinetics use components not listed in the reactants... $\rightarrow$
incomplete PKN (missing parents)

# Conclusion

- ▶ SBML2BN = pipeline for the automatic transformation of quantitative SBML models into compatible Boolean Networks
- ▶ Overall, the pipeline is good (runtime & quality of the Boolean networks synthesised)
- ▶ Possible improvements for even more complex models (models with $> 10$ parents).

Thanks for your attention.
Hope to {see, read} you.
athenais.vaginay@loria.fr

Enjoy the conference!

# Biomodels repository



1039 manually curated models (data from 08th Nov. 2021)

- Ordinary differential equation model
- Logical model
- Constraint-based model
- Petri net