

Teatime capsid:

Automatic transformation from reaction models to Boolean models using answer-sets constrained by a topology and an abstracted dynamic.

Athénaïs Vaginay

2020 january 23rd (enfin \o/ :D)

Teatime capsid:

Automatic transformation from **reaction models** to **Boolean models** using **answer-sets** constrained by a topology and an abstracted dynamic.

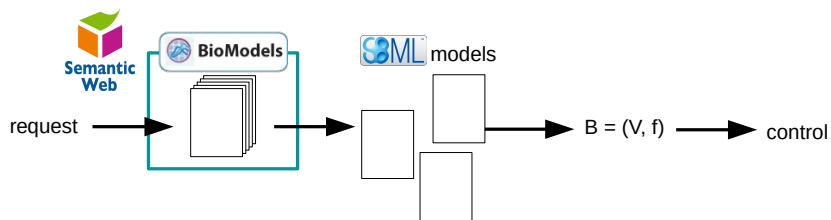
Athénaïs Vaginay

2020 january 23rd (enfin \o/ :D)

what did we talked about one year ago ?
(time flies...)

Project Thesis Overview

(which i like sooooo much ♡)



as automatized as possible... ;-)

Modèle : définition du Dictionnaire de l'Académie française

Modèle, n.m., XVI^e siècle, *modelle*. Emprunté de l'italien *modello*, de même sens.

4. Sciences. **Représentation** physique, graphique ou, plus généralement, **mathématique qui formalise les relations unissant les différents éléments d'un système**, d'un processus, d'une structure, **en vue de faciliter la compréhension de certains mécanismes ou de permettre la validation d'une hypothèse**. [...]



(To find your way around a city, you do not need to have a map at scale 1:1...)

Biological levels

Many levels in biology → multi scale biological models.

vertical hierarchy

biosphere
ecosystem
community
population
organism
organ
tissue
→ cell ←
macromolecule
molecule
atom

horizontal modularity

(signaling, gene regulation, metabolism)

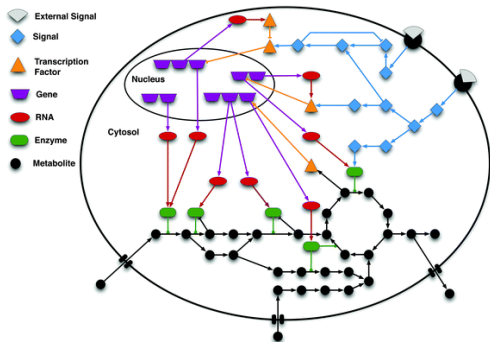


Figure: Gonçalves et al. "Bridging the Layers: Towards Integration of Signal Transduction, Regulation and Metabolism into Mathematical Models." *Molecular BioSystems* 9, no. 7 (2013): 1576. <https://doi.org/10.1039/c3mb25489e>

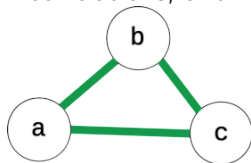
modelling framework

a LOT of formalisms

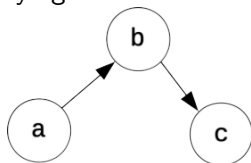
- ▶ statistical model
- ▶ mechanistic models

formalisms — statistical models

based on correlations, undirectional.



for an underlying mechanism that looks like:

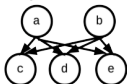


formalisms — mechanistic models

Boolean network



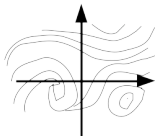
Bayesian network



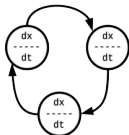
Process algebras

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

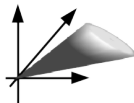
Differential equation



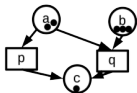
Hybrid systems



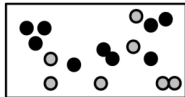
Constraint based model



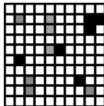
Petri Nets



Agent-based model



Cellular automata



Interacting state machine

Compartment based

Rule based

...

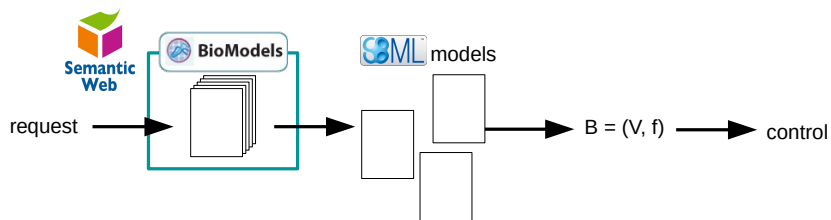
community standards

standards needed to allow model sharing (and tools to communicate)



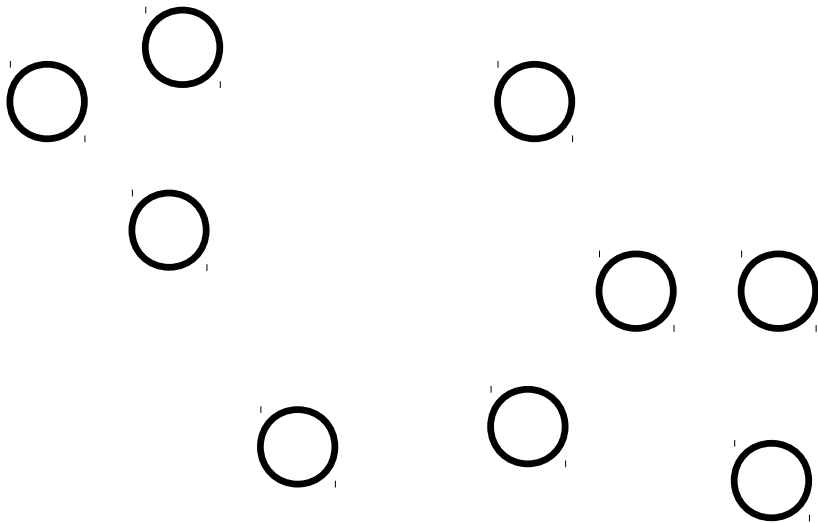
SBML, based on reaction model, is agnostic to the formalism.

Project Thesis Overview (sometimes I wonder if I'll have time $\neg_(_)_/_$)

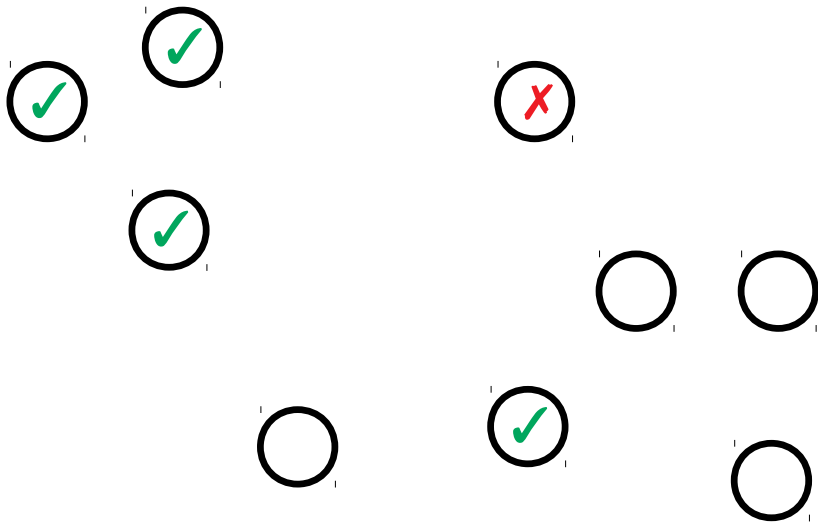


as automatized as possible... ;-)

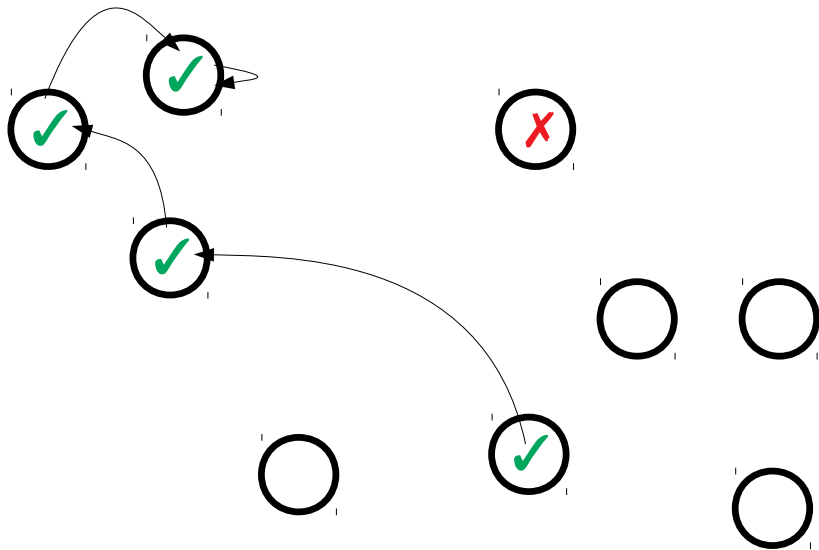
About control — the possible states of a system



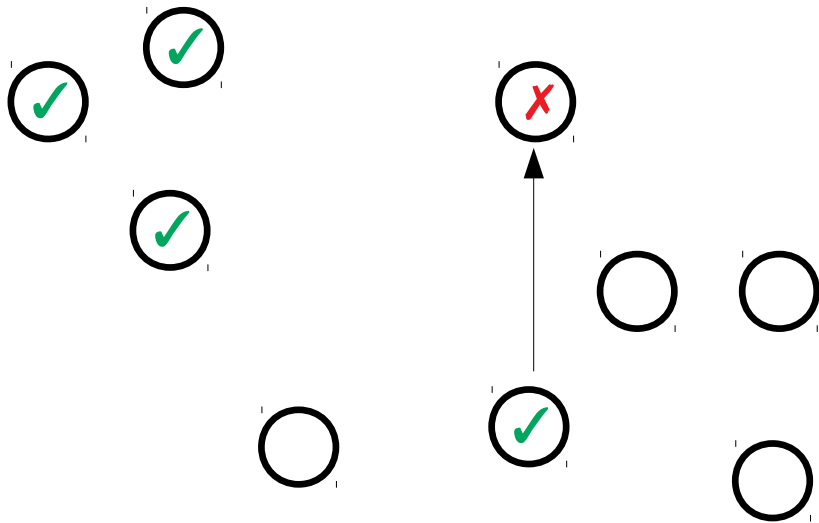
About control — some are “normal”, some are ‘pathological’



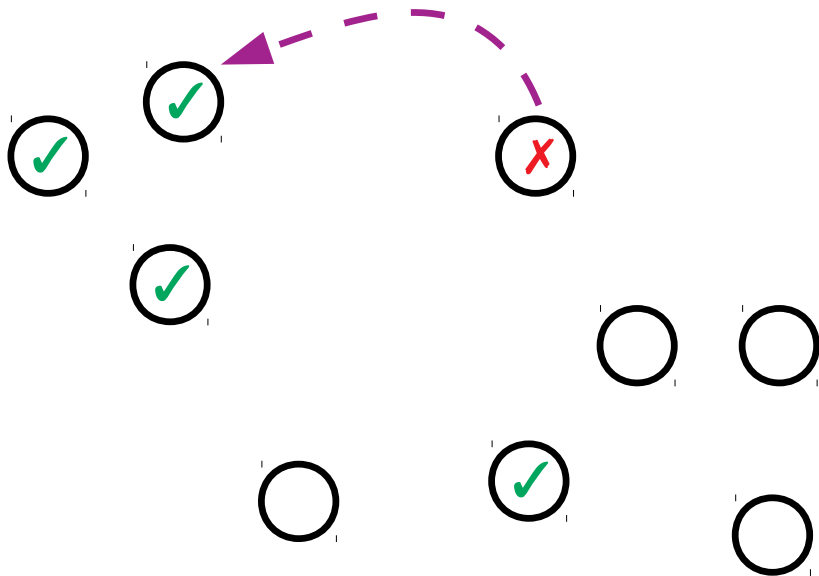
About control — behaviour of the **normal** system



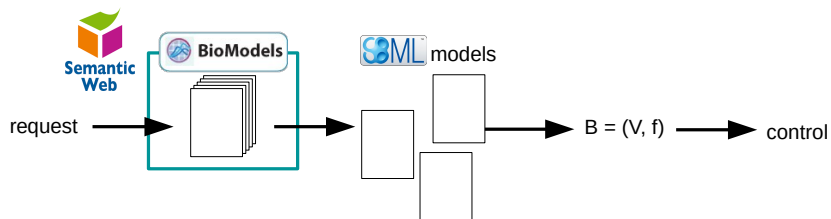
About control — behavior of the **broken** system



About control — behavior of the **repaired** system



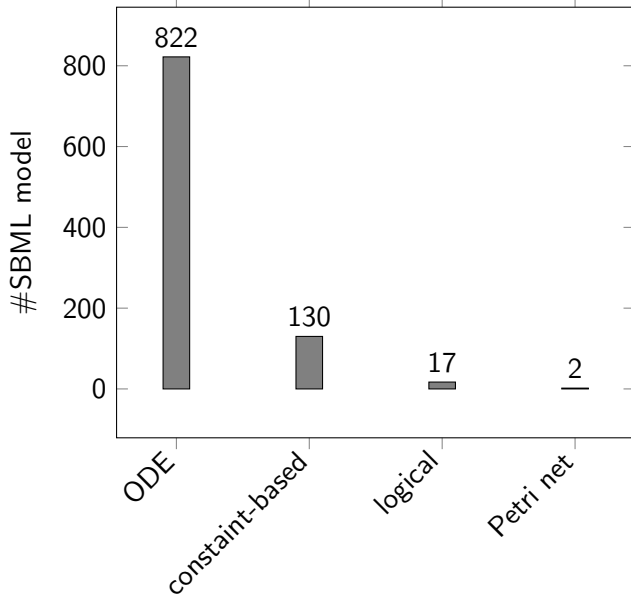
Project Thesis Overview



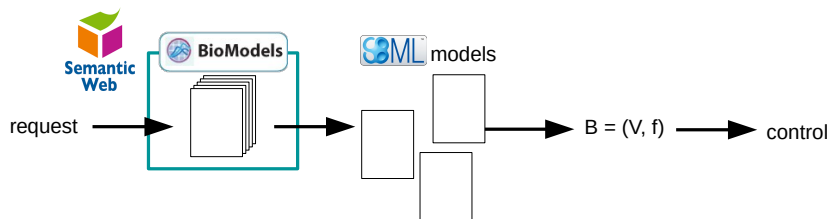
as automatized as possible... ;-)

BioModels content (updated 2019 dec 12)

1943 SBML models at total. 831 curated.



Project Thesis Overview



as automatized as possible... ;-)

Dish of the day (“Et bon appétit, bien sûr”)

 model



$B = (V, f)$

as automatized as possible... ;-)



Steps overview

- ▶ what is a boolean network. Concretely...? (easy-ish)
- ▶ retrieve the list of components (super archi easy)
- ▶ retrieve the influences between component (easy : use the Theorems of Fages 2008)
- ▶ retrieve the binarized behavior of the system (easy : just solve deterministically the equa diff then apply a binarization threshold)
- ▶ overapproximate the dynamic using the reachability (mediumly hard...)
- ▶ encode this in ASP (OMG, haaaaaaaaaard)

Boolean network formalism = network of Boolean automata

$$\mathbb{B} = \{0, 1\}$$

$$B = f : \mathbb{B}^n \rightarrow \mathbb{B}^n$$

$$B = (V, f_i : \mathbb{B}^n \rightarrow \mathbb{B} \forall i \in V)$$

boolean network formalism = network of boolean automata

- ▶ automata
- ▶ automata configuration
- ▶ system configuration
- ▶ influence
- ▶ fonctions locales de transitions
- ▶ updating scheme
- ▶ transition graph
- ▶ reachability



Automates

© Musée des arts et métiers-Cnam
Sylvain Pelly

Warning

any resemblance to elements of fiction already existing **can not** be only fortuitous. . .

Les Feux de l'Amour

A stylized red calligraphic logo consisting of the letters 'F' and 'A' intertwined. The 'F' is at the top, with a long horizontal stroke that loops under the 'A'. The 'A' is below the 'F', with a long vertical stroke that loops back up to the 'F'. The overall style is elegant and artistic.

3 automata: $P, N, J \in V$

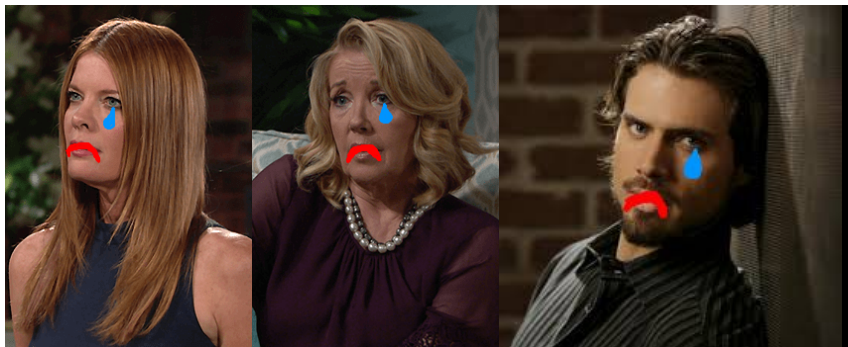
- ▶ Phillis (P) girlfriend of Joshua
- ▶ Nikki (N) is the mother of Joshua and does not like Phillis
- ▶ Joshua (J) is the son of Nikki and boyfriend of Phillis



Boolean automata \rightarrow their state is Boolean

P, N and J can be either sad (0) or happy (1).

Automata state: boolean $\forall i \in V, x_i \in \mathbb{B} = \{0, 1\}$



Boolean automata \rightarrow their state is Boolean

P, N and J can be either sad (0) or happy (1).

Automata state: boolean $\forall i \in V, x_i \in \mathbb{B} = \{0, 1\}$



system configuration : $x \in \mathbb{B}^n$

Here $n = 3 \rightarrow 2^3 = 8$ different configurations.

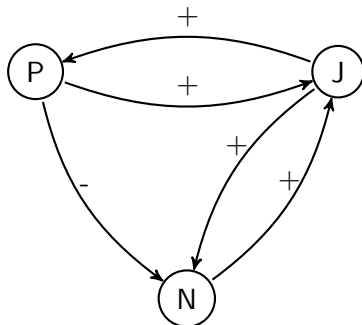
P	N	J	
0	0	0	0
0	0	1	1
0	1	0	2
0	1	1	3
1	0	0	4
1	0	1	5
1	1	0	6
1	1	1	7

$\Delta(x, y) := \{i \in \{1, \dots, n\} \mid x_i \neq y_i\}$: set of component which differ between two configurations x and y .

des relations tout à fait classiques. . .

- ▶ P loves J. $IN(P) = \{J\}$, $k_P = 1$
- ▶ N loves her son J but not P $IN(N) = \{J, P\}$, $k_N = 2$
- ▶ J loves P and his dear mom N (differently, but still. . .)
 $IN(J) = \{P, N\}$, $k_J = 2$

influence graph:

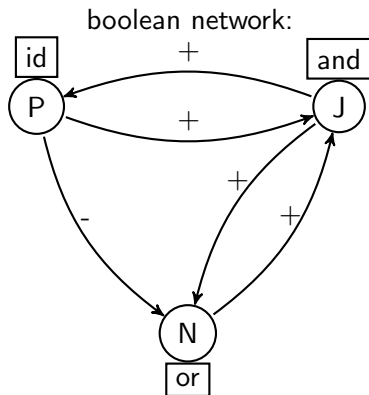


$IN(x_i) = \{x_{i_1}, \dots, x_{i_k}\}$: the set of k parents nodes of x_i .

k : the “indegree” of x_i . K : the maximum indegree of a BN.

des dynamiques tout à fait classiques...

- ▶ P happy if J happy: $f_P = x_J$
- ▶ N happy if J happy or if P not happy: $f_N = x_J \vee \neg x_P$
- ▶ J happy when both N and P happy: $f_J = x_P \wedge x_N$



fonction locale de transition : $f_{x_i} : \mathbb{B}^{k_i} \rightarrow \mathbb{B}$

the truth (table) of the crisis

$$f_P = x_J$$

$$f_N = x_J \vee \neg x_P$$

$$f_J = x_P \wedge x_K$$

x_P	x_N	x_J	$f_P(x)$	$f_N(x)$	$f_J(x)$
0	0	0	0	1	0
0	0	1	1	1	0
0	1	0	0	1	0
0	1	1	1	1	0
1	0	0	0	0	0
1	0	1	1	1	0
1	1	0	0	0	1
1	1	1	1	1	1

transition graph

The transition graph describes the behavior of the system.

$$\mathcal{G} = (V_{\mathcal{G}} = \mathbb{B}^n, E_{\mathcal{G}} \subseteq \mathbb{B}^n \times \mathbb{B}^n)$$

It depends on the updating scheme.

updating scheme = organisation of the updates in time

Any order or any number of automata can be updated at each time step \rightarrow infinite number of possibility.

Some classical:

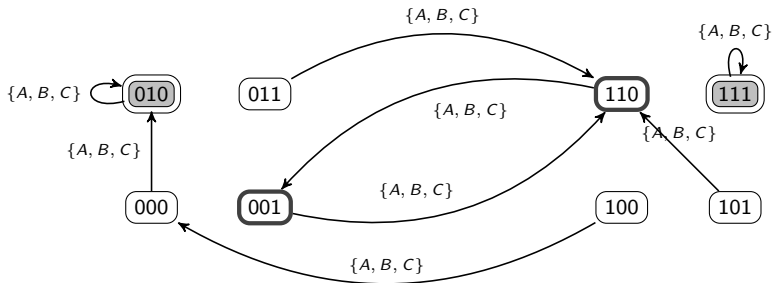
- ▶ (deterministic) synchronous (or parallele)
- ▶ (pure) asynchronous
- ▶ generalized asynchronous (or elementary)
- ▶ deterministic sequential block
- ▶ ...

One weird that we will use:

- ▶ most permissive (in $\mathbb{M} = \{0, *, 1\}$)

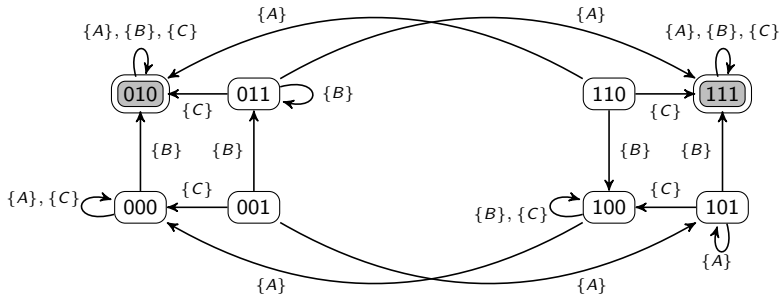
synchronous $\mathcal{G}^\pi : \{A, B, C\}$

all components get updated at each time step : $x \xrightarrow{\pi} y$ iff $f(x) = y$



true asynchronous $\mathcal{G}^\alpha : \{\{A\}, \{B\}, \{C\}\}$

only one component (chosen non deterministically) is updated at each time step: $x \xrightarrow{\alpha} y$ iff $\exists i \in V, \Delta(x, y) = i \wedge y_i = f_i(x)$



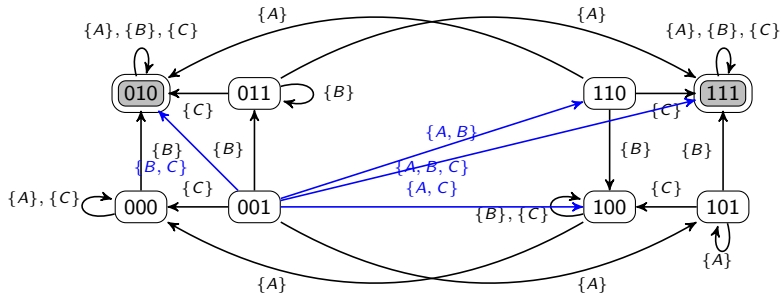
$$f_P = x_J$$

$$f_N = x_J \vee \neg x_P$$

$$f_J = x_P \wedge x_K$$

generalized asynchronous $\mathcal{G}_{010}^\epsilon : \{\{A\}, \{B\}, \{C\}, \{A, B\}, \{A, C\}, \{B, C\}, \{A, B, C\}\}$

any number of component get updated non deterministically in one time step:
 $x \xrightarrow{\epsilon} y$ iff $\forall i \in \Delta(x, y), y_i = f_i(x)$



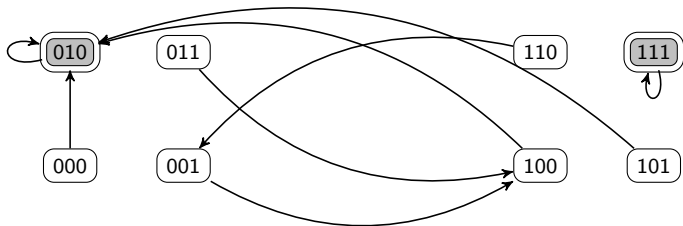
$$f_P = x_J$$

$$f_N = x_J \vee \neg x_P$$

$$f_J = x_P \wedge x_K$$

sequential blocs $\mathcal{G}^{\beta_1} = (\{A, B\}, \{C\})$

updates following a given block order

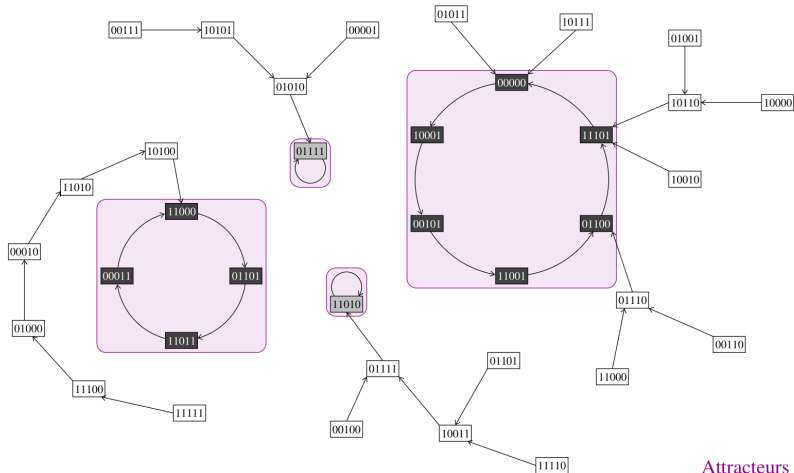


$$f_P = x_J$$

$$f_N = x_J \vee \neg x_P$$

$$f_J = x_P \wedge x_K$$

Trajectory and attractors (cycles and fixed points)



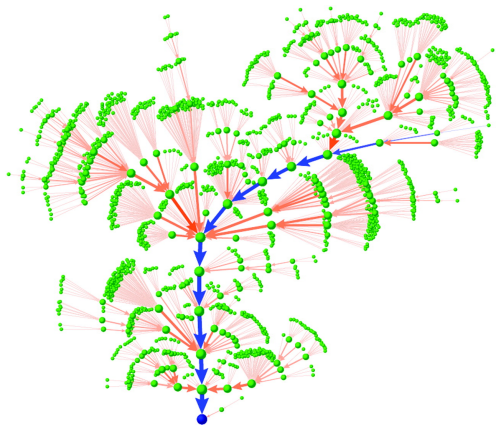
Attracteurs

Image from Sylvain Sené "Réseaux d'automates et systèmes biologiques une approche par le calcul naturel", Biorégul 2019

Trajectory and attractors — biology mapping example

Any boolean model has at least 1 attractor.

Attractors are mapped to (biological) properties (they represent physiological functions, cellular types, ...).



Dynamical trajectories of the 1,764 protein states (green nodes) flowing to the G1 fixed point (blue node). Arrows between states indicate the direction of dynamic flow from one state to another. The cell-cycle sequence is colored blue. The size of a node and the thickness of an arrow are proportional to the logarithm of the traffic flow passing through them.

From "The yeast cell-cycle network is robustly designed." Li et al. 2004. (Fig. 2)

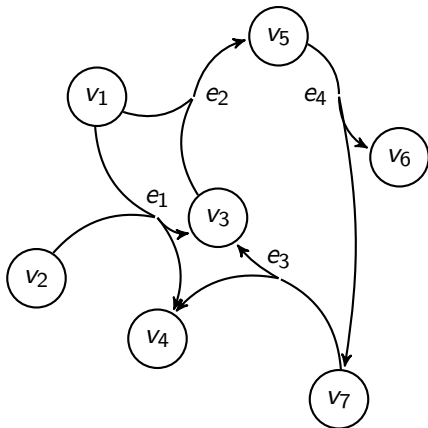
reachability

The fact that there exist a sequence with which a configuration is accessible from another configuration.

Steps overview

- ▶ what is a boolean network. Concretely...? (easy-ish)
- ▶ retrieve the list of components (super archi easy)
- ▶ retrieve the influences between component (easy : use the Theorems of Fages 2008)
- ▶ retrieve the binarized behavior of the system (easy : just solve deterministically the equa diff then apply a binarization threshold)
- ▶ overapproximate the dynamic using the reachability (mediumly hard...)
- ▶ encode this in ASP (OMG, haaaaaaaaaard)

$$M = \{e_i \text{ for } R_i = [M_i] \Rightarrow P_i\}_{i=1\dots n}$$



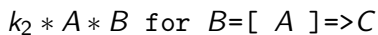
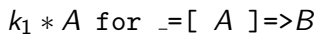
hypergraph = nodes + hyperedges

Differential semantic for reaction models

$$M = \{e_i \text{ for } R_i = [M_i] \Rightarrow P_i\}_{i=1\dots n}$$

$$dx/dt = \sum_{i=1}^n (P_i(x) - R_i(x)) * e_i$$

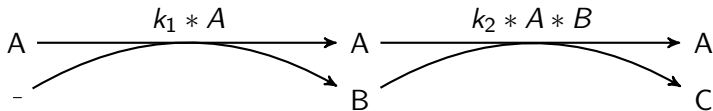
an example :



$$dA/dt = 0$$

$$dB/dt = k_1 * A - k_2 * A * B$$

$$dC/dt = k_2 * A * B$$



Steps overview

- ▶ what is a boolean network. Concretely...? (easy-ish)
- ▶ retrieve the list of components (super archi easy)
- ▶ retrieve the influences between component (easy : use the Theorems of Fages 2008)
- ▶ retrieve the binarized behavior of the system (easy : just solve deterministically the equa diff then apply a binarization threshold)
- ▶ overapproximate the dynamic using the reachability (mediumly hard...)
- ▶ encode this in ASP (OMG, haaaaaaaaaard)

retrieve the influence graph from the SBML file

influence graph: a directed graph with labeled edges
(+: positive influence; -: negative influence).

Theorem from Fages et al. 2008 prove the equivalence between Reaction Models and Influence Graph. Either:

- ▶ Parse the rules, and determine the influence from the stoichiometric coefficients given in the rules.
- ▶ Use the jacobian matrix J of the ODE model formed of the partial derivatives $J_{i,j} = \frac{\partial i}{\partial j}$.

Steps overview

- ▶ what is a boolean network. Concretely...? (easy-ish)
- ▶ retrieve the list of components (super archi easy)
- ▶ retrieve the influences between component (easy : use the Theorems of Fages 2008)
- ▶ retrieve the binarized behavior of the system (easy : just solve deterministically the equa diff then apply a binarization threshold)
- ▶ overapproximate the dynamic using the reachability (mediumly hard...)
- ▶ encode this in ASP (OMG, haaaaaaaaaard)

Solve the equa diff deterministically

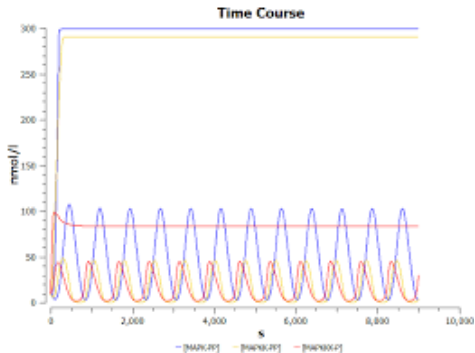


Figure: from “Building and Simulating Models using COPASI”, © 2016, Nicolas LeNovère, Viji Chelliah, Bhupinder Virk. creative commons Attribution - Share Alike 4.0 licence.

Scale between 0 and 1, then apply a threshold

ρ_t^x : continuous observation $\in [0; 1]$ of a component x at time t ,
 β_t^x : its Boolean value ; determined with the following :

$$\beta_t^x \triangleq \begin{cases} 1 & \text{when } \rho_t^x \leq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

Other methods exist, but this one is the most simple one.

Steps overview

- ▶ what is a boolean network. Concretely...? (easy-ish)
- ▶ retrieve the list of components (super archi easy)
- ▶ retrieve the influences between component (easy : use the Theorems of Fages 2008)
- ▶ retrieve the binarized behavior of the system (easy : just solve deterministically the equa diff then apply a binarization threshold)
- ▶ overapproximate the dynamic using the reachability (mediumly hard...)
- ▶ encode this in ASP (OMG, haaaaaaaaaard)

meta-configuration in most permissive semantics

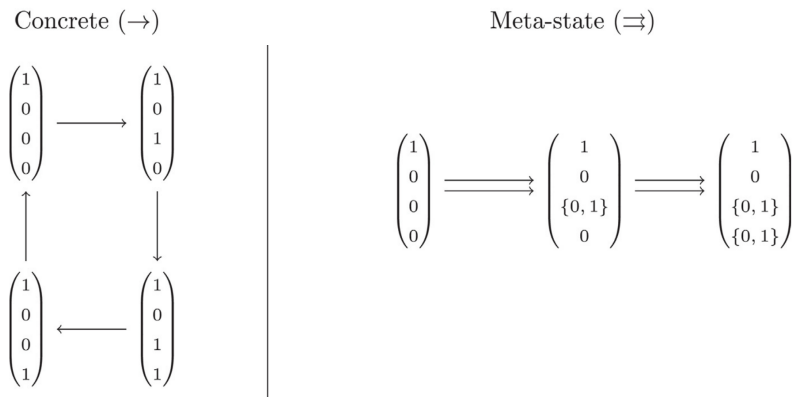
Reachability : the fact that there exist a sequence with which a configuration is accessible from another configuration.

Most permissive semantics consists in **approximating** the dynamic using the **reachability** between states and not directly the state transition.

A meta-configuration concatenates together all the configurations that are reachable from a given configuration.

*One Semantic to approximate them all,
One Semantic to find them,
One Semantic to bring them all
and in a vector bind them.*

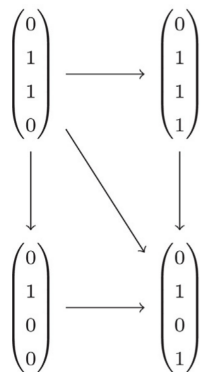
meta-configurations : example 1



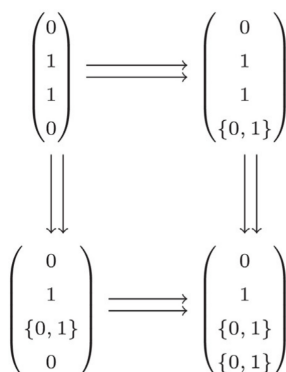
cycle of the BN transition graph is over-approximated by a fixed-point in the meta-state semantics

meta-configurations : example 2

Concrete (\rightarrow)



Meta-state (\Rightarrow)



fixed-point of the BN transition graph is over-approximated by a fixed-point in the meta-state semantics

Steps overview

- ▶ what is a boolean network. Concretely...? (easy-ish)
- ▶ retrieve the list of components (super archi easy)
- ▶ retrieve the influences between component (easy : use the Theorems of Fages 2008)
- ▶ retrieve the binarized behavior of the system (easy : just solve deterministically the equa diff then apply a binarization threshold)
- ▶ overapproximate the dynamic using the reachability (mediumly hard...)
- ▶ encode this in ASP (OMG, haaaaaaaaaard)

ASP = Answer Set Programming

- ▶ declarative language
- ▶ rules are written in the form of clauses:
H :- B1; ...; Bn.
- ▶ that correspond to logical implications:
H if B1 and ... and Bn.
- ▶ “generate then filter” approach
- ▶ a solution is a True / False instantiation of atoms that respect the rules
- ▶ ASP program can generate “answer sets” = the set of all the solutions that would work

Crazy hard to sum the element of a table. But 6 lines are needed to solve a sudoku...

sum the elements of a table in ASP

```
val(1;3;18;23;42;67;72).
```

```
first(X):- val(X) ; not val(Y): val(Y), Y<X.
```

```
next(X,Y):- val(X) ; val(Y) ; X<Y ; not val(Z): val(Z), X<Z, Z<Y. ■
```

```
sum(1,F,F):- first(F).
```

```
sum(N+1,W,T+W):- sum(N,V,T) ; next(V,W).
```

```
step(S):- sum(S,-,-).
```

```
sum(S):- sum(N,-,S) ; not sum(M,-,-): step(M), N<M.
```

```
#show sum/1. % Pouloulou !
```

Conclusion : “on en bave des ronds de chapeau”

sudoku solving in ASP

```
x(1..9). y(1..9). n(1..9).
```

```
{sudoku(X,Y,N): n(N)}=1 :- x(X);y(Y).
```

```
subgrid(X,Y,A,B) :- x(X); x(A); y(Y); y(B);  
                    (X-1)/3==(A-1)/3; (Y-1)/3==(B-1)/3.
```

```
:- sudoku(X,Y,N); sudoku(A,Y,N); X!=A.
```

```
:- sudoku(X,Y,N); sudoku(X,B,N); Y!=B.
```

```
:- sudoku(X,Y,V); sudoku(A,B,V); subgrid(X,Y,A,B); X != A; Y != B.
```

```
#show sudoku/3.
```

Our problem in ASP = TODO

Ask ASP what are the logical rules of a given topology and that behave such that their dynamic fullfill the constraints about the reachability between the configurations?

For the next tea time... ;-)

Thanks for your attention...
Any questions ? :)

References I

Extra slides

.

sum element of a table in ASP

```
% Values to be sum up :
val(1;3;18;23;42;67;72).
% Find the first one :
first(X):- val(X) ; not val(Y): val(Y), Y<X.
% Find the "next" one following an increasing order
next(X,Y):- val(X) ; val(Y) ; X<Y ; not val(Z): val(Z),
% Add the thing in the list , one after the other
sum(1,F,F):- first(F). % At first step , this corres. to
sum(N+1,W,T+W):- sum(N,V,T) ; next(V,W). % At step n, i
step(S):- sum(S,-,-).
% The total is the last sum at the last step
sum(S):- sum(N,-,S) ; not sum(M,-,-): step(M), N<M.
#show sum/1. % Pouloulou !
```

Conclusion : "on en bave des ronds de chapeau"

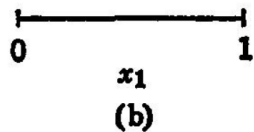
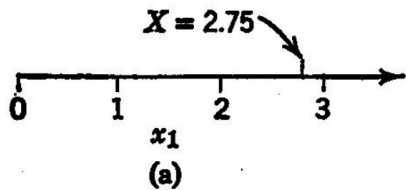
influence graph — Fages 2008

influence graph: a directed graph with labeled edges (+: positive influence; -: negative influence).

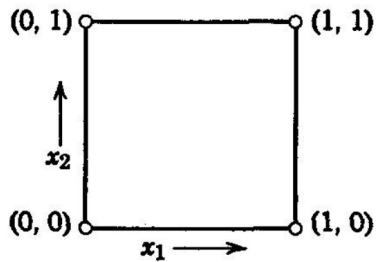
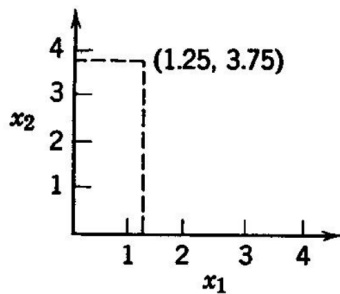
Formally, 2 definitions (Fages et al. 2008):

- ▶ Synthactical Influence Graph (SIG), from the stoichiometric coefficients given in the rules.
- ▶ Differential Influence Graph (DIG), from the jacobian matrix J of the ODE model formed of the partial derivatives $J_{i,j} = \frac{\partial x}{\partial j}$.

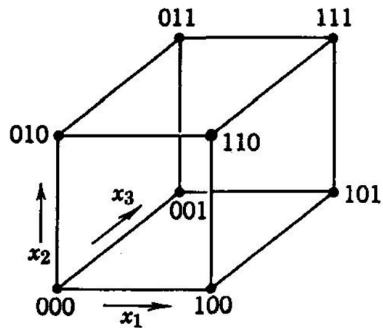
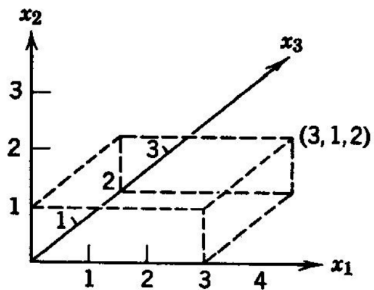
hypercubical representation



hypercubical representation

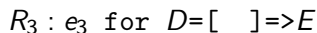
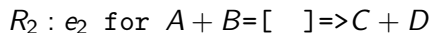
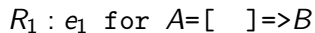


hypercubical representation



Reaction model

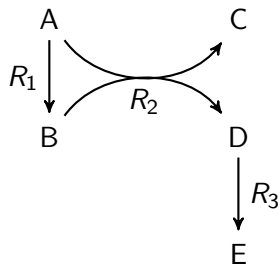
Reactions:



Stoichiometric matrix:

$$\begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} \begin{pmatrix} R_1 & R_2 & R_3 \\ -1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Hypergraph:



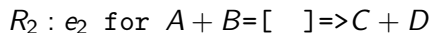
citation: Klamt <https://journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable>

[//journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable](https://journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable)

citation Fages 2018

Reaction model with stoichiometry

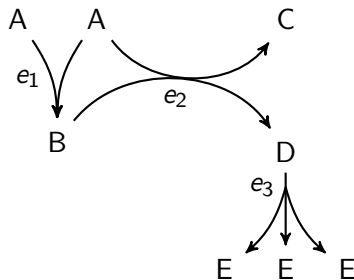
Reactions:



Stoichiometric matrix:

$$\begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} \begin{pmatrix} R_1 & R_2 & R_3 \\ -2 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix}$$

Hypergraph:



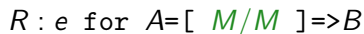
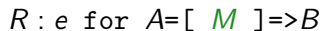
citation: Klamt <https://journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable>

[//journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable](https://journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable)

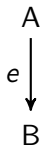
citation Fages 2018

Reaction model with modifier

Reaction:



Hypergraph :



Stoichiometric matrix:

$$\begin{matrix} & R \\ A & \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \\ B & \\ M & \end{matrix}$$

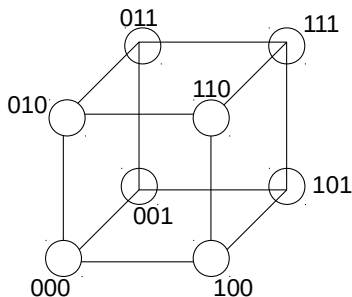
citation : Klamt <https://journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable>

<https://journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1000385&type=printable>
citation Fages 2018

meta-states are hypercubes

set of associated configurations :

$$c(h) = \{x \in \mathbb{B}^n \mid \forall i \in \{1, \dots, n\}, h_i \neq * \Rightarrow x_i = h_i\}$$



Meta-configurations in meta semantics

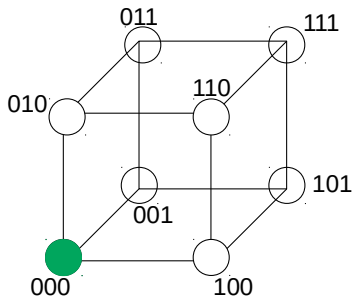
\mathbb{M}^n



meta-states are hypercubes

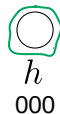
set of associated configurations :

$$c(h) = \{x \in \mathbb{B}^n \mid \forall i \in \{1, \dots, n\}, h_i \neq * \Rightarrow x_i = h_i\}$$



Configurations in
concrete semantic
 \mathbb{B}^n

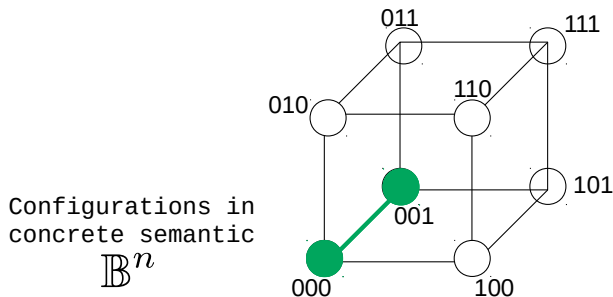
Meta-configurations in
meta semantics
 \mathbb{M}^n



meta-states are hypercubes

set of associated configurations :

$$c(h) = \{x \in \mathbb{B}^n \mid \forall i \in \{1, \dots, n\}, h_i \neq * \Rightarrow x_i = h_i\}$$



Meta-configurations in meta semantics

M^n

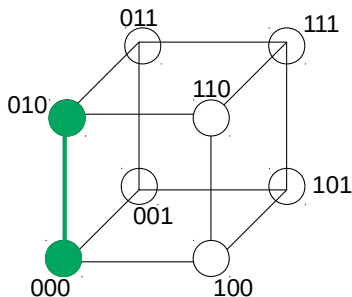


h
 00^*

meta-states are hypercubes

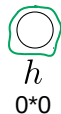
set of associated configurations :

$$c(h) = \{x \in \mathbb{B}^n \mid \forall i \in \{1, \dots, n\}, h_i \neq * \Rightarrow x_i = h_i\}$$



Configurations in
concrete semantic
 \mathbb{B}^n

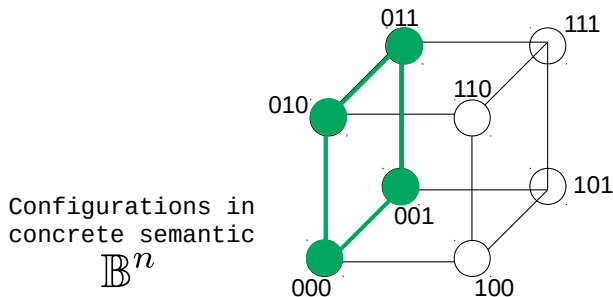
Meta-configurations in
meta semantics
 \mathbb{M}^n



meta-states are hypercubes

set of associated configurations :

$$c(h) = \{x \in \mathbb{B}^n \mid \forall i \in \{1, \dots, n\}, h_i \neq * \Rightarrow x_i = h_i\}$$



Meta-configurations in meta semantics

M^n

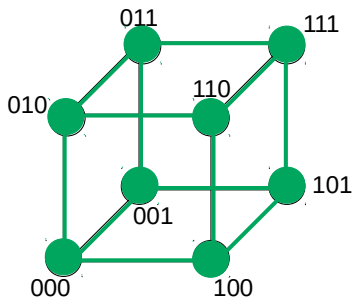


h
 0^{**}

meta-states are hypercubes

set of associated configurations :

$$c(h) = \{x \in \mathbb{B}^n \mid \forall i \in \{1, \dots, n\}, h_i \neq * \Rightarrow x_i = h_i\}$$



Meta-configurations in meta semantics

M^n



h
