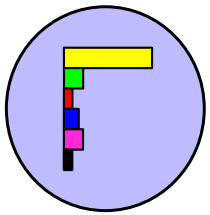
# Méthodes formelles pour l'étude de la dynamique des réseaux biologiques

Loïc Paulevé

CNRS/LRI, Univ. Paris-Sud, Univ. Paris-Saclay – BioInfo team
loic.pauleve@lri.fr
http://loicpauleve.name

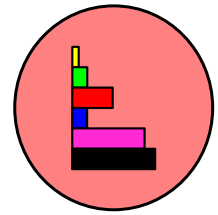
Journées GdR BiM - 14 Novembre 2017

# Cellular Dynamics



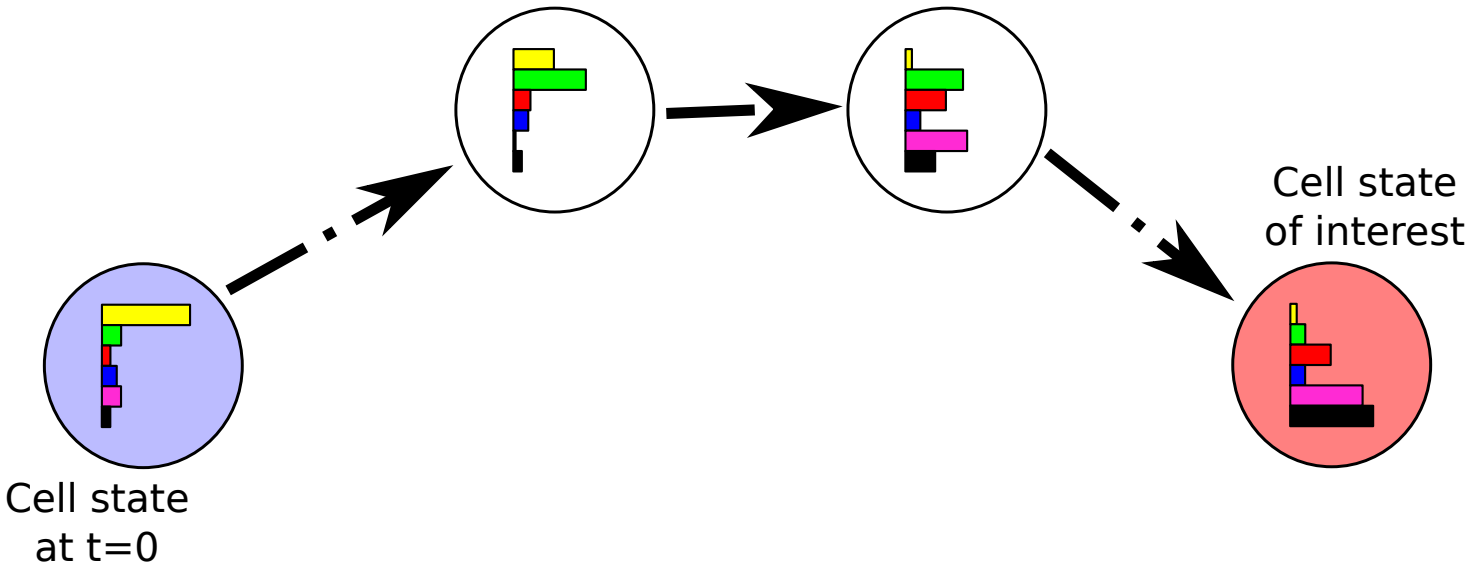
Cell state
at t=0

Cell state
of interest



Initial state(s)/Goal state(s)

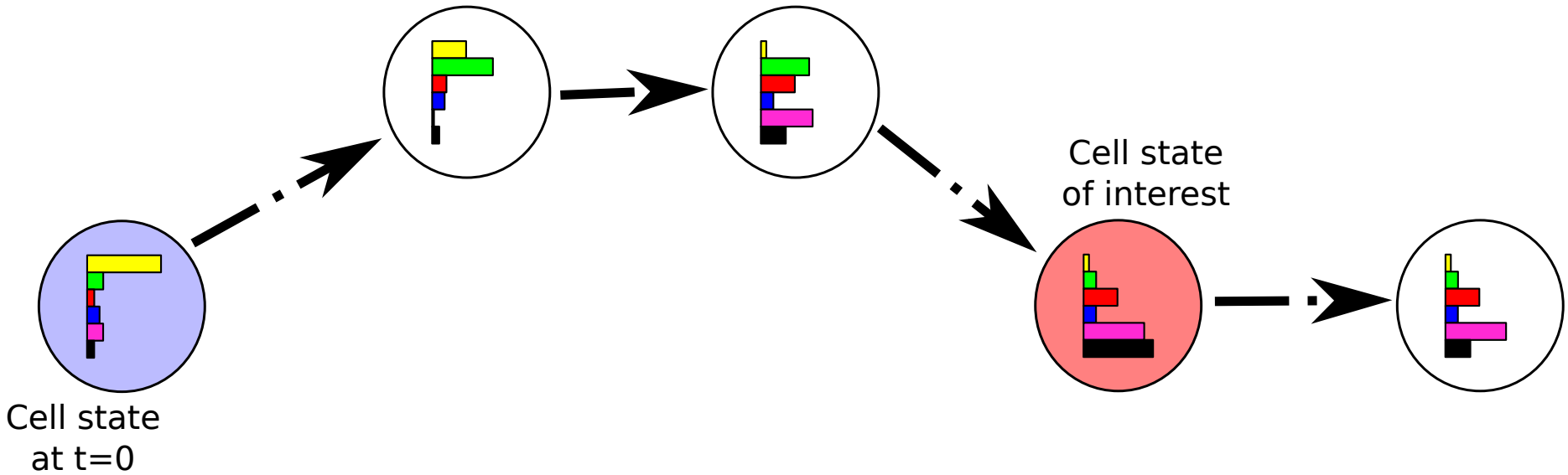
## Cellular Dynamics




Initial state(s)/Goal state(s)

- Trajectory existence (reachability)

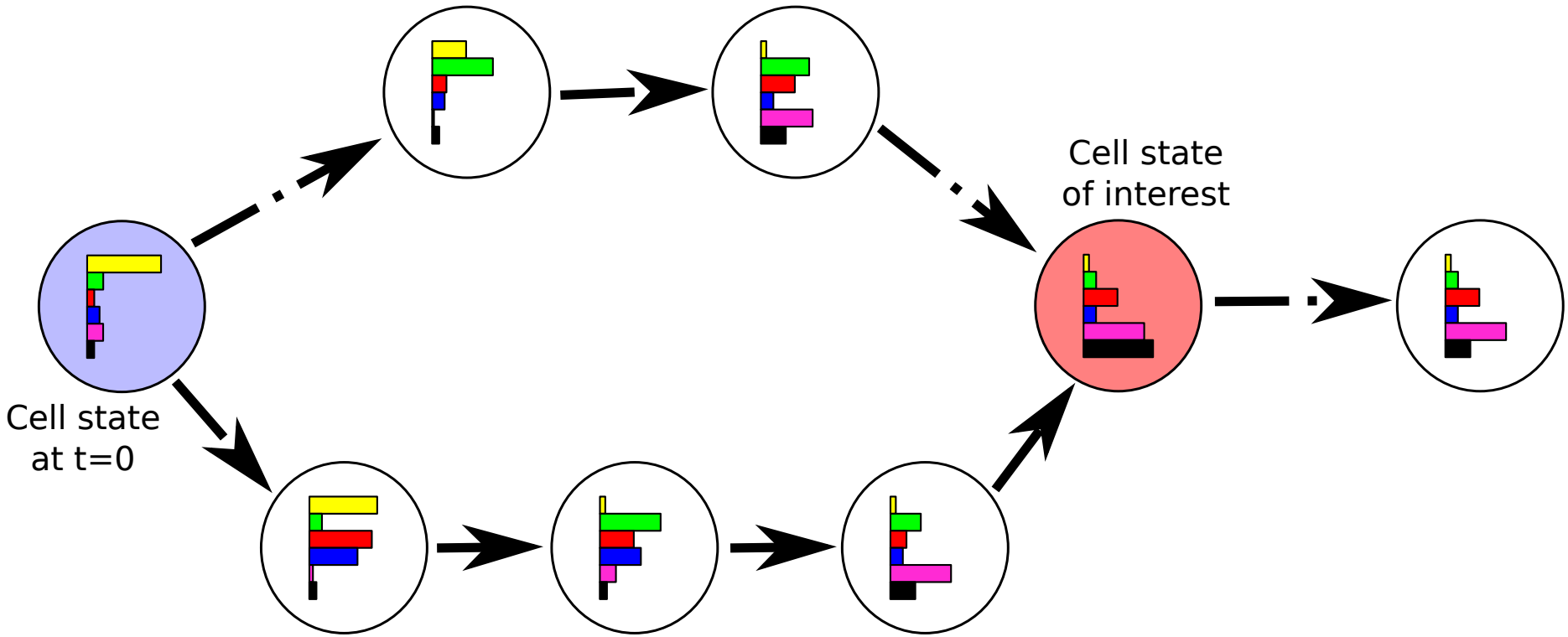
## Cellular Dynamics




Initial state(s)/Goal state(s)

- Trajectory existence (reachability)

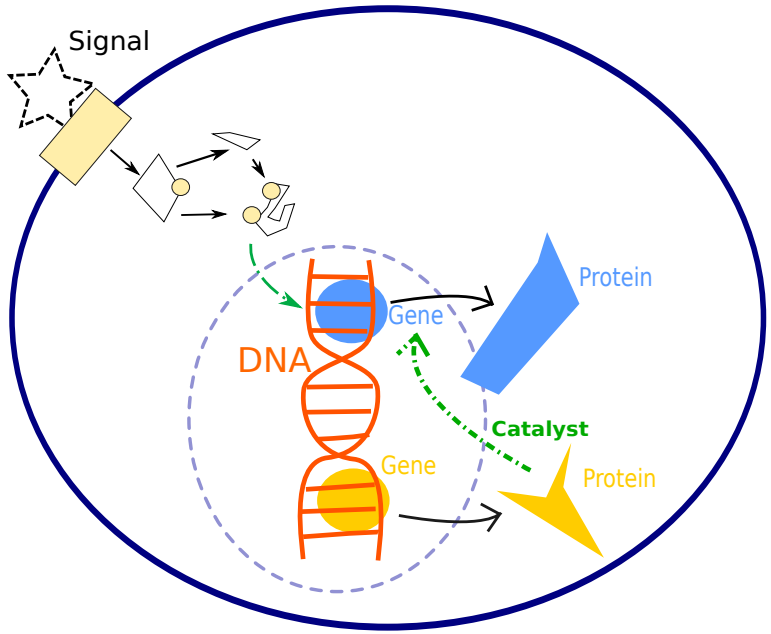
# Cellular Dynamics




Initial state(s)/Goal state(s)

- Trajectory existence (reachability)
- Reasoning on all trajectories : e.g., common features
- **Control** : perturbations to avoid/enforce goal reachability

# Signalling and gene networks



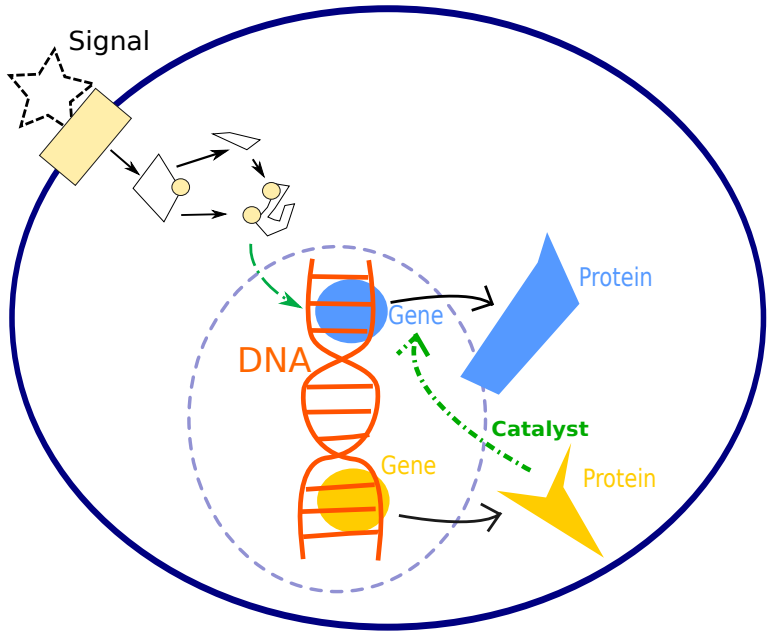

**Prediction**

- Cell response w.r.t. signal+environment
- Long-term behaviours (differentiation)

**Control**

- Mutations/Perturbations for modifying cell behaviour
- Trans/De-differentiation

# Signalling and gene networks




**Prediction**

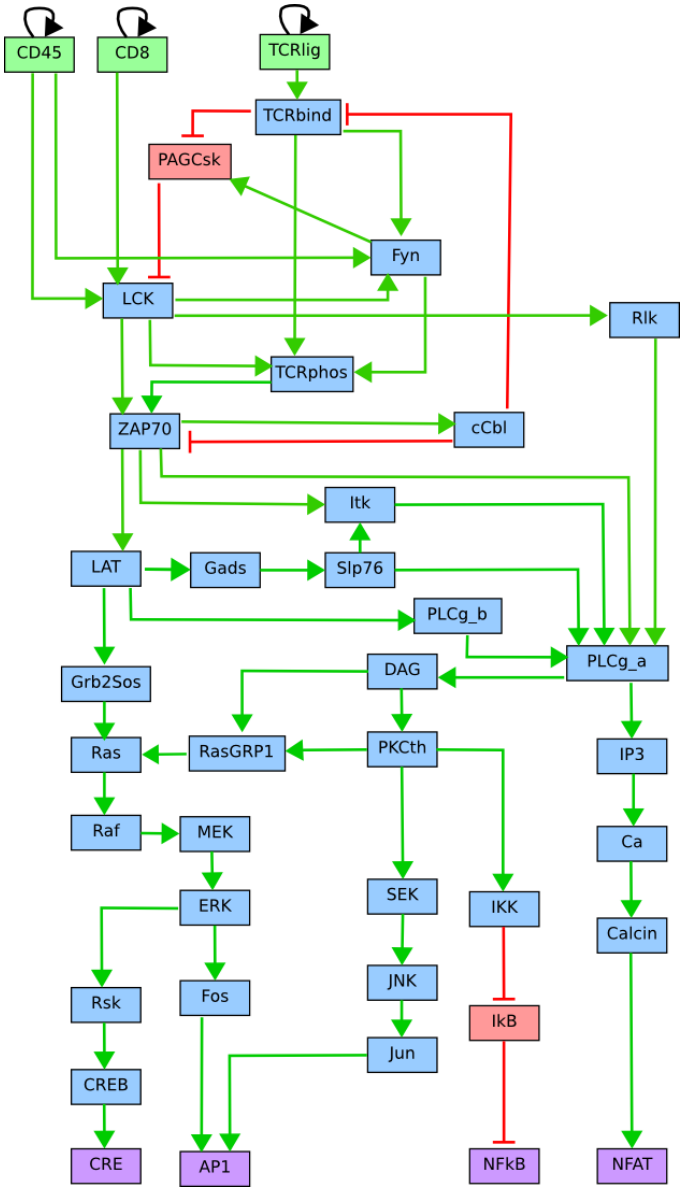
- Cell response w.r.t. signal+environment
- Long-term behaviours (differentiation)

**Control**

- Mutations/Perturbations for modifying cell behaviour
- Trans/De-differentiation

$$\Rightarrow \begin{cases} \textbf{Computational models of dynamics} \\ -\text{Formal verification} \\ -\text{Automatic reasoning} \end{cases}$$

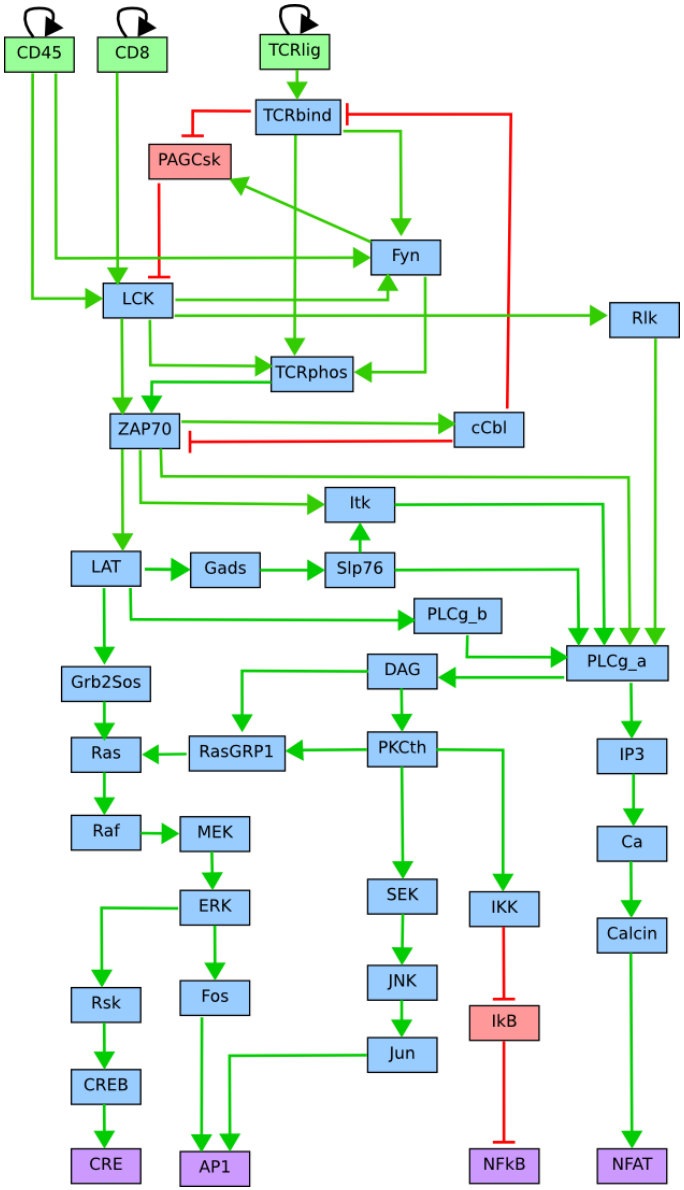
# Computational models of biological networks




[Naldi et al, PLOS Comput Biol 2010]

Network : account for **indirect influences** between **entities** of a system

# Computational models of biological networks



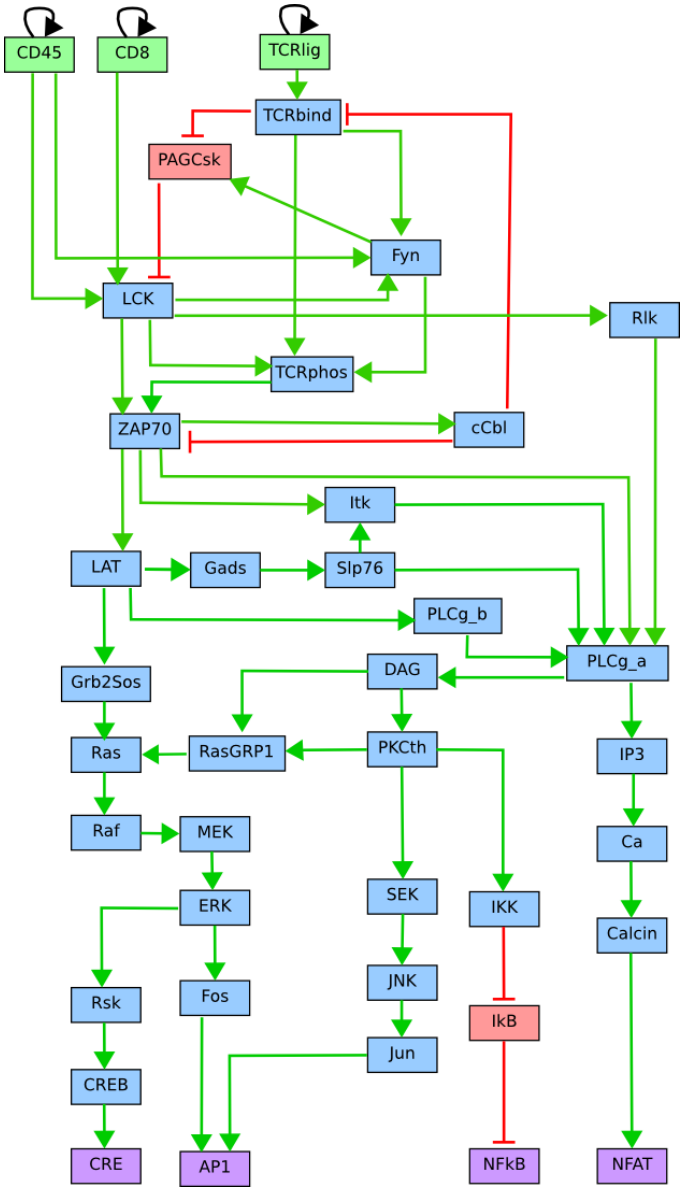

[Naldi et al, PLOS Comput Biol 2010]

Network : account for **indirect influences** between **entities** of a system

A biological model is typically built from

- literature (tedious)
- (curated) databases : pull interactions discovered in very different experimental settings
- network inference from data : prune networks to fit with data ; identify new interactions
- expert knowledge (people)

# Computational models of biological networks




[Naldi et al, PLOS Comput Biol 2010]

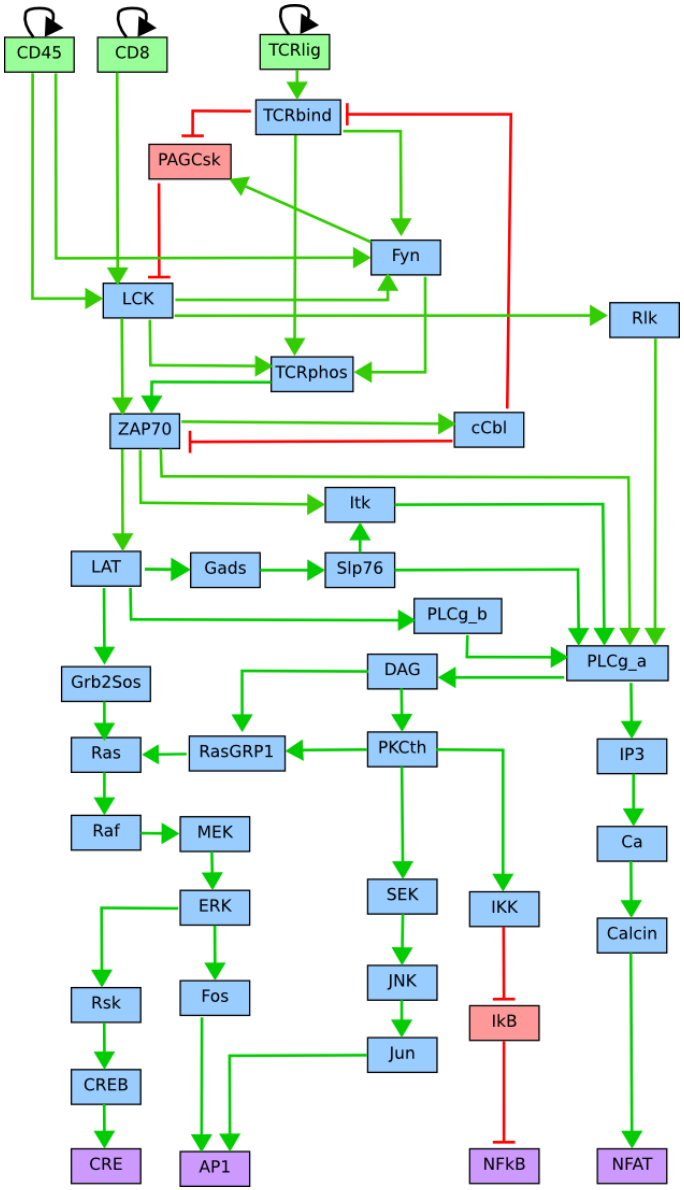
Network : account for **indirect influences** between **entities** of a system

A biological model is typically built from

- literature (tedious)
- (curated) databases : pull interactions discovered in very different experimental settings
- network inference from data : prune networks to fit with data ; identify new interactions
- expert knowledge (people)

⇒ uncertainties / hypotheses remain
⇒ set of candidate models

# Computational models of biological networks




[Naldi et al, PLOS Comput Biol 2010]

Network : account for **indirect influences** between **entities** of a system

A biological model is typically built from

- literature (tedious)
- (curated) databases : pull interactions discovered in very different experimental settings
- network inference from data : prune networks to fit with data ; identify new interactions
- expert knowledge (people)

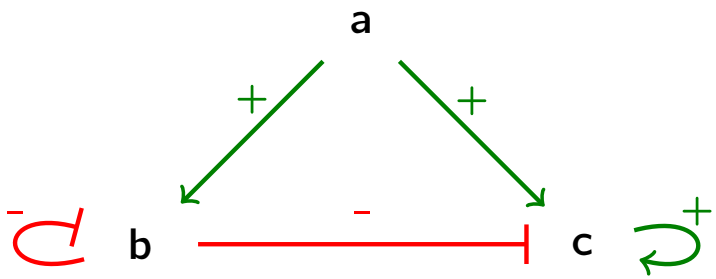
⇒ uncertainties / hypotheses remain
⇒ set of candidate models

Need for efficient methods to

- discriminate, refine candidate models
- predictions robust to model uncertainties.

# Computational models of biological networks

## Knowledge




## + Semantics

Ordinary differential equations

$$\frac{da}{dt} = -k_{da}a$$

$$\frac{db}{dt} = \frac{k_{ab}a}{1+k_{ab}a}\frac{1}{1+k_{bb}b} - k_{db}b$$

$$\frac{dc}{dt} = \left(\frac{k_{ac}a}{1+k_{ac}a}\frac{k_{cc}c}{1+k_{cc}c}+\right)\frac{1}{1+k_{bc}b} - k_{dc}c$$

Boolean network

$f_a(a, b, c) = 0$

$f_b(a, b, c) = a$ and not $b$

$f_c(a, b, c) =$ not $b$ and ($a$ or $c$)

**Semantics**

- Mathematically defines what a state is,
- and how it evolves with time (sequences or chronometry)
- Requires additional parameters, usually not in knowledge

## Motivating question

Given a computational model of a network,
how to prove that a behaviour is impossible ?

Example : it is impossible to reach the state of interest in the current condition

This question is key for :

- **Model verification** : do we miss something ?
- **Model identification** : filter valid candidate models
- **Control prediction** : perturbations which makes a behaviour impossible

## Motivating question

Given a computational model of a network,
how to prove that a behaviour is impossible ?

Example : it is impossible to reach the state of interest in the current condition

This question is key for :

- **Model verification** : do we miss something ?
- **Model identification** : filter valid candidate models
- **Control prediction** : perturbations which makes a behaviour impossible

Requires a complete assessment of model capabilities :

- ~~simulation~~
- **formal verification**

## Motivating question

Given a computational model of a network,
how to prove that a behaviour is impossible ?

Example : it is impossible to reach the state of interest in the current condition

This question is key for :

- **Model verification** : do we miss something ?
- **Model identification** : filter valid candidate models
- **Control prediction** : perturbations which makes a behaviour impossible

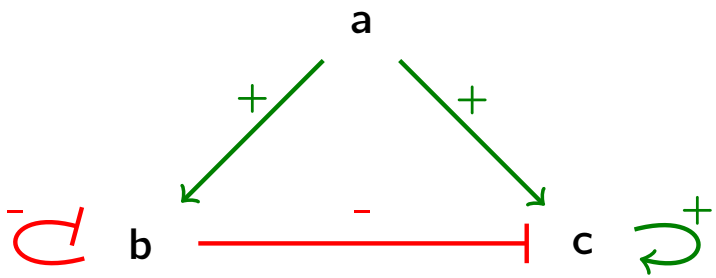
Requires a complete assessment of model capabilities :

- ~~simulation~~
- **formal verification**

Same principle to prove absence of bugs in computer programs
$\Rightarrow$ similar technologies, different models.

# Dynamics of Qualitative Networks

## Example in Boolean case




$f_a(a, b, c) = 0$

$f_b(a, b, c) = a$ and not $b$

$f_c(a, b, c) =$ not $b$ and ($a$ or $c$)

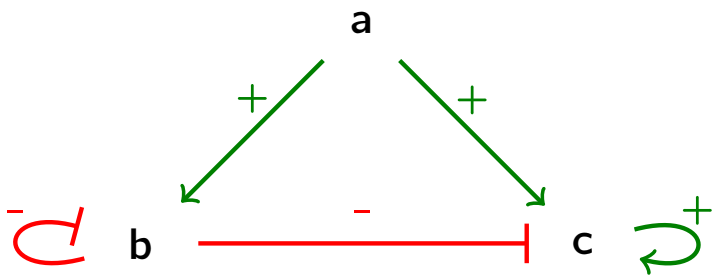
**State transition graph**

$\langle a, b, c \rangle$ $\langle 1, 0, 0 \rangle$

[René Thomas in Journal of Theoritical Biology, 1973] [A. Richard, J.-P. Comet, G. Bernot in Modern Formal Methods and Applications, 2006]

# Dynamics of Qualitative Networks

## Example in Boolean case

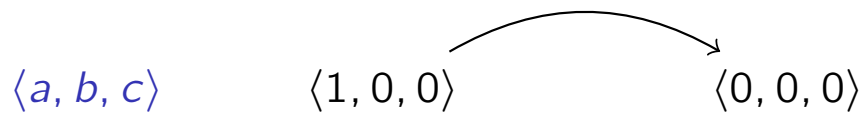



$f_a(a, b, c) = 0$

$f_b(a, b, c) = a$ and not $b$

$f_c(a, b, c) =$ not $b$ and ($a$ or $c$)

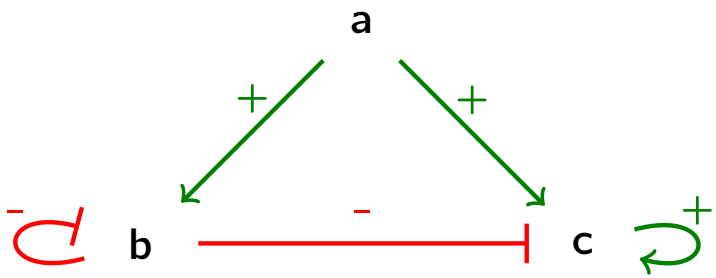
**State transition graph**




[René Thomas in Journal of Theoritical Biology, 1973] [A. Richard, J.-P. Comet, G. Bernot in Modern Formal Methods and Applications, 2006]

# Dynamics of Qualitative Networks

## Example in Boolean case

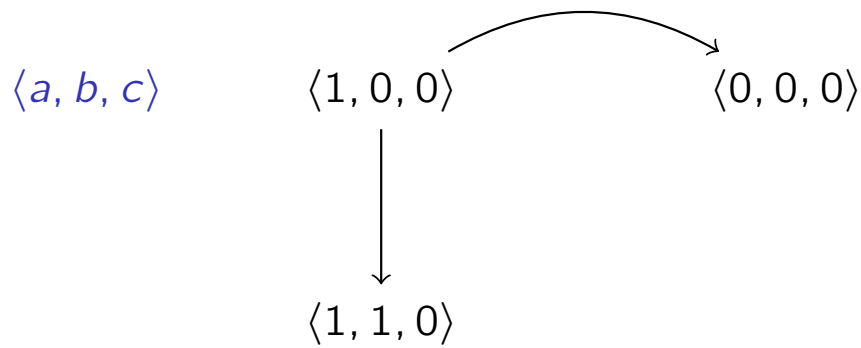



$f_a(a, b, c) = 0$

$f_b(a, b, c) = a$ and not $b$

$f_c(a, b, c) =$ not $b$ and ($a$ or $c$)

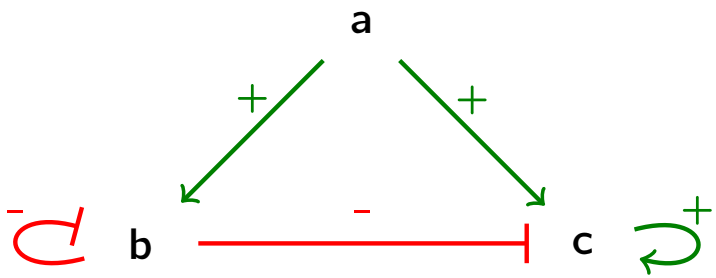
**State transition graph**




[René Thomas in Journal of Theoritical Biology, 1973] [A. Richard, J.-P. Comet, G. Bernot in Modern Formal Methods and Applications, 2006]

# Dynamics of Qualitative Networks

## Example in Boolean case

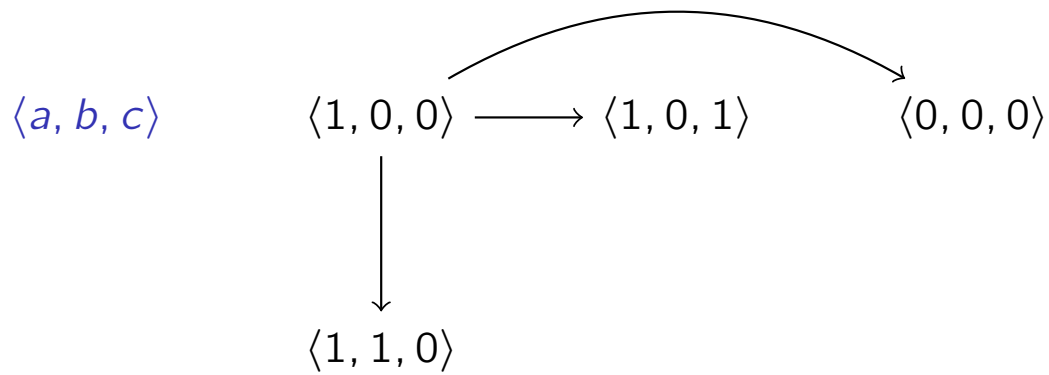



$f_a(a, b, c) = 0$

$f_b(a, b, c) = a$ and not $b$

$f_c(a, b, c) = \text{not } b \text{ and } (a \text{ or } c)$

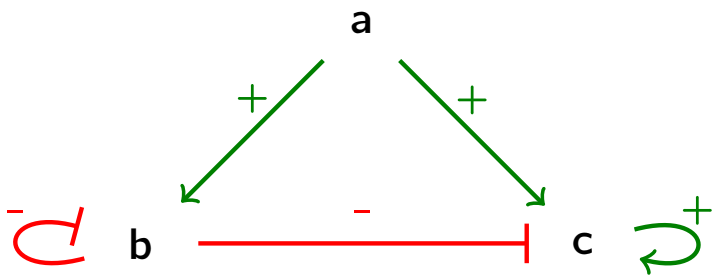
**State transition graph**




[René Thomas in Journal of Theoritical Biology, 1973] [A. Richard, J.-P. Comet, G. Bernot in Modern Formal Methods and Applications, 2006]

# Dynamics of Qualitative Networks

## Example in Boolean case

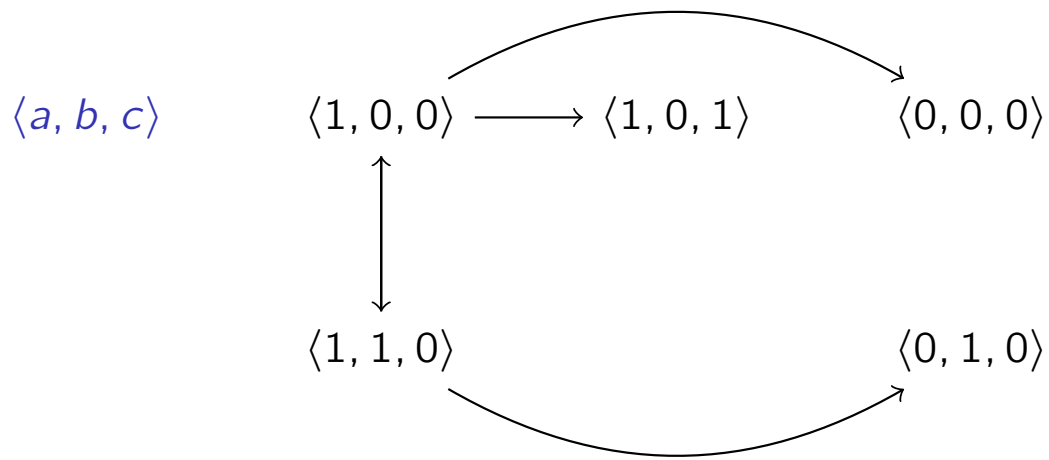



$f_a(a, b, c) = 0$

$f_b(a, b, c) = a$ and not $b$

$f_c(a, b, c) = \text{not } b \text{ and } (a \text{ or } c)$

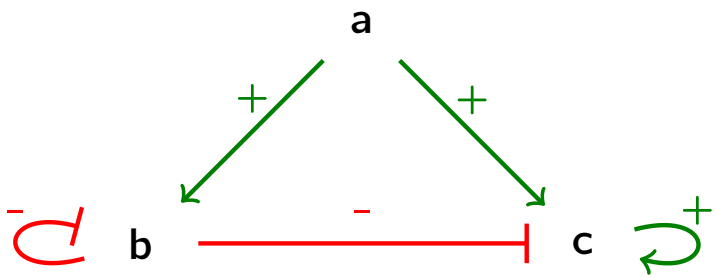
**State transition graph**




[René Thomas in Journal of Theoritical Biology, 1973] [A. Richard, J.-P. Comet, G. Bernot in Modern Formal Methods and Applications, 2006]

# Dynamics of Qualitative Networks

## Example in Boolean case

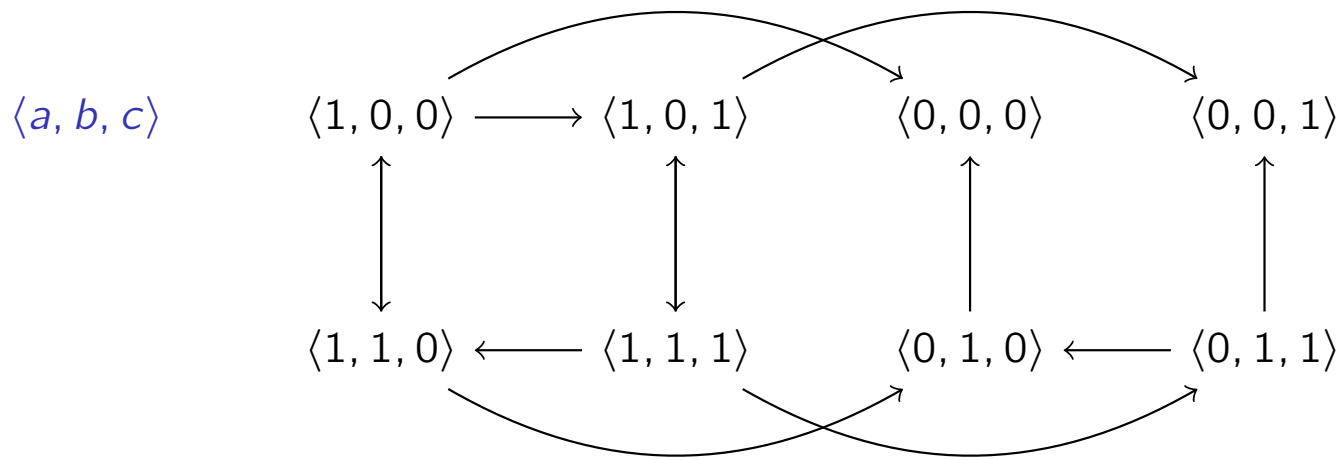



$f_a(a, b, c) = 0$

$f_b(a, b, c) = a$ and not $b$

$f_c(a, b, c) = \text{not } b \text{ and } (a \text{ or } c)$

**State transition graph**




[René Thomas in Journal of Theoritical Biology, 1973] [A. Richard, J.-P. Comet, G. Bernot in Modern Formal Methods and Applications, 2006]

# Formal Verification of Qualitative Networks

**Qualitative models**

- Focus on causality of state changes
- Few parameters :
  $\Rightarrow$ quite direct translation from knowledge to computational model
  $\Rightarrow$ results are general (independant of speed of reactions, precise quantities. . . )
- Not suited for quantitative predictions.

# Formal Verification of Qualitative Networks

**Qualitative models**

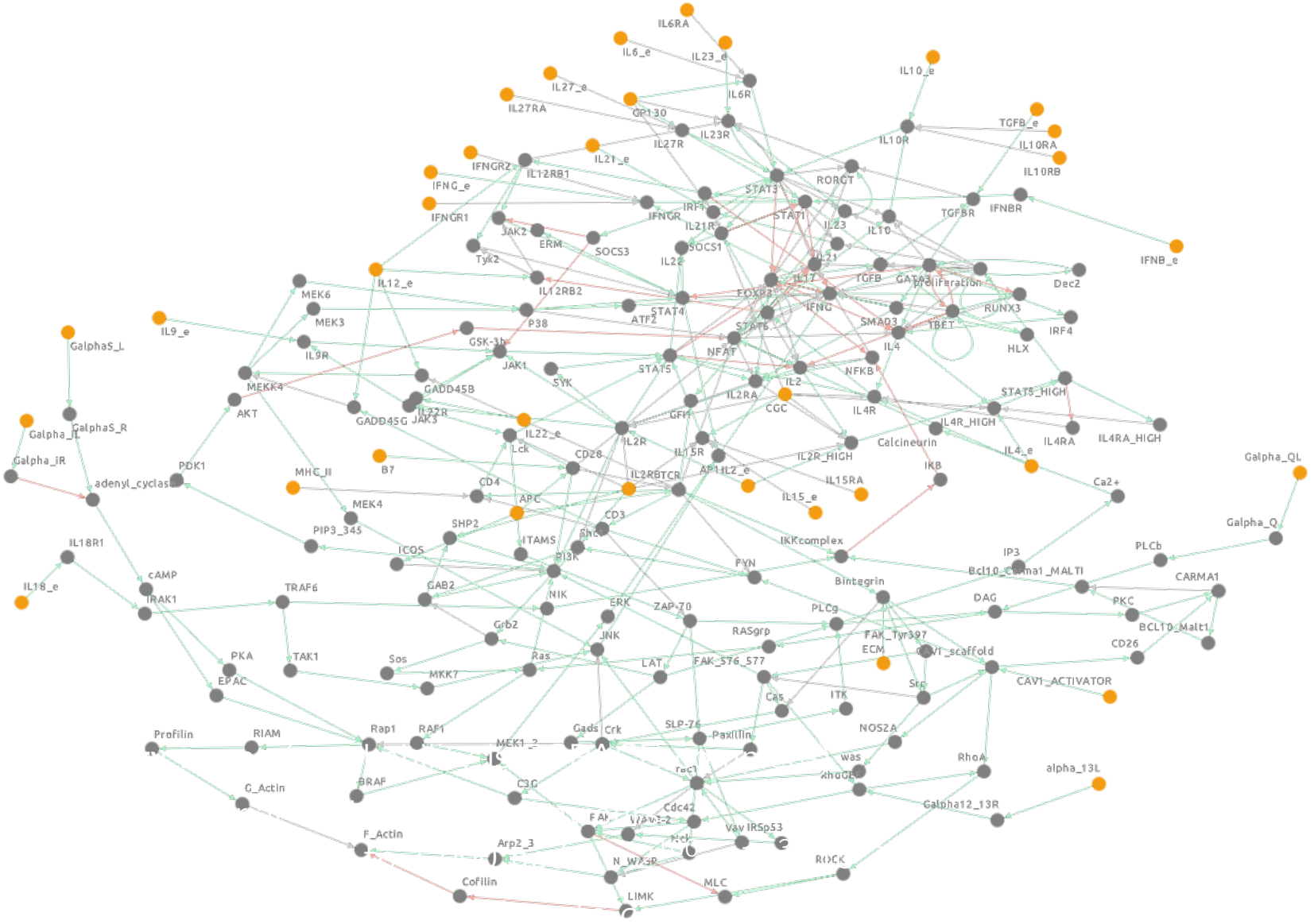
- Focus on causality of state changes
- Few parameters :
  ⇒ quite direct translation from knowledge to computational model
  ⇒ results are general (independant of speed of reactions, precise quantities. . . )
- Not suited for quantitative predictions.

**State transition graph**

- Allows an exhaustive view of model capabilities ;
- Automatic "model checking" w.r.t. specifications.

but. . .

## Formal Verification of Qualitative Networks




[Source: https://cellcollective.org/#3521/cd4-t-cell-signaling]

# Formal Verification of Qualitative Networks

**Qualitative models**

- Focus on causality of state changes
- Few parameters :
  $\Rightarrow$ quite direct translation from knowledge to computational model
  $\Rightarrow$ results are general (independant of speed of reactions, precise quantities. . . )
- Not suited for quantitative predictions.

**State transition graph**

- Allows an exhaustive view of model capabilities ;
- Automatic "model checking" w.r.t. specifications.

but. . .

**Tractability issues** (reachability is PSPACE-complete)

- Combinatorial explosion of behaviours
  networks with 100 to 1,000 nodes : $2^{100}$ - $10^{30}$ to $2^{1\,000}$ - $10^{300}$ states
- Large range of initial conditions to consider.
- Difficult to extract comprehensive proofs of (im)possibility.

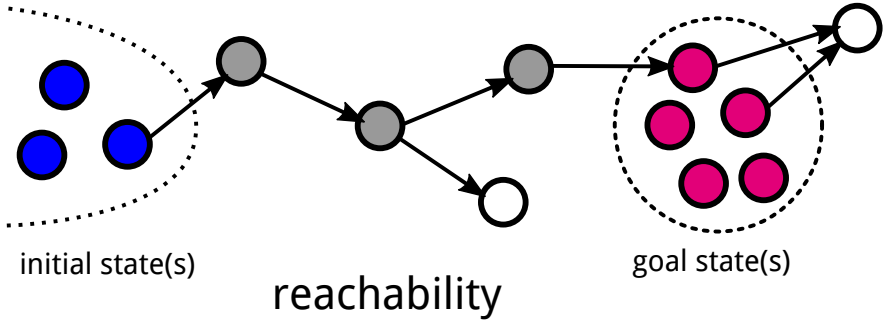
## Abstractions for transient dynamics of Boolean Networks

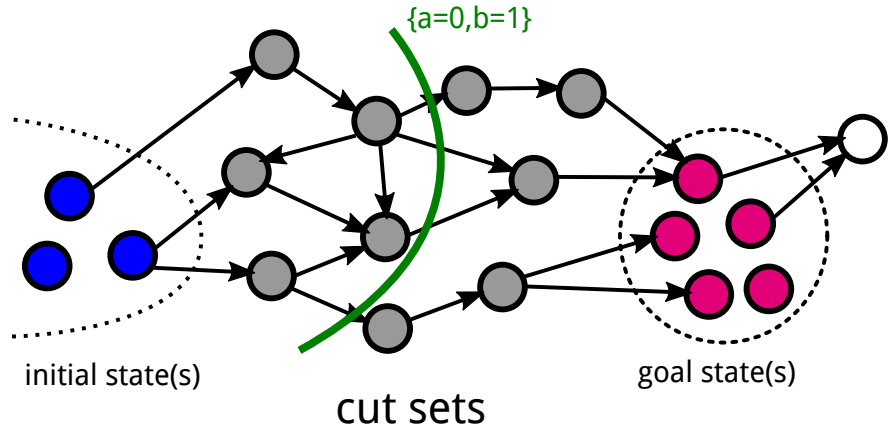
*Intuition* : exploit the low scope of transitions (concurrency)

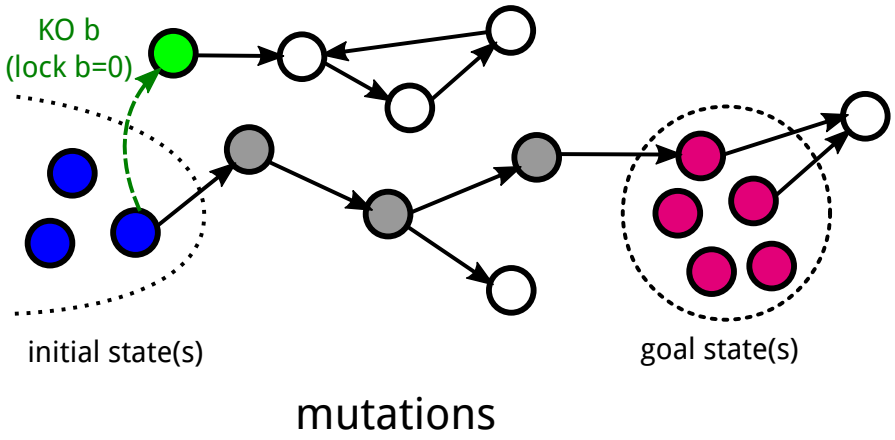
- Static analysis by abstract interpretation [Cousot and Cousot 77]
- Intermediate representation (Local Causality Graph)
- Gives necessary/sufficient conditions for reachability
- Implementation in SAT/Answer-Set Programming (ASP)

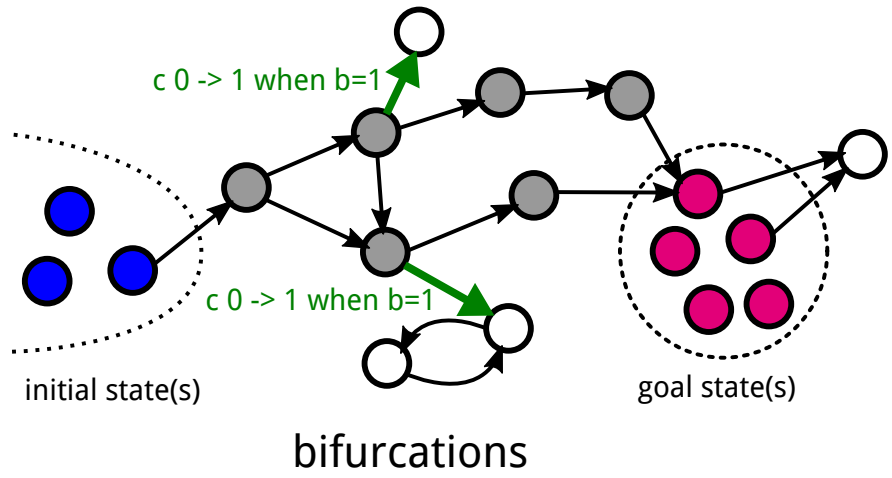
$\Rightarrow$ **Approx. of PSPACE problems with P.$e^d$ or NP.$e^d$ problems**
where $d$ is the in-degree of nodes in the Boolean network

# Causal analysis for transient reachability



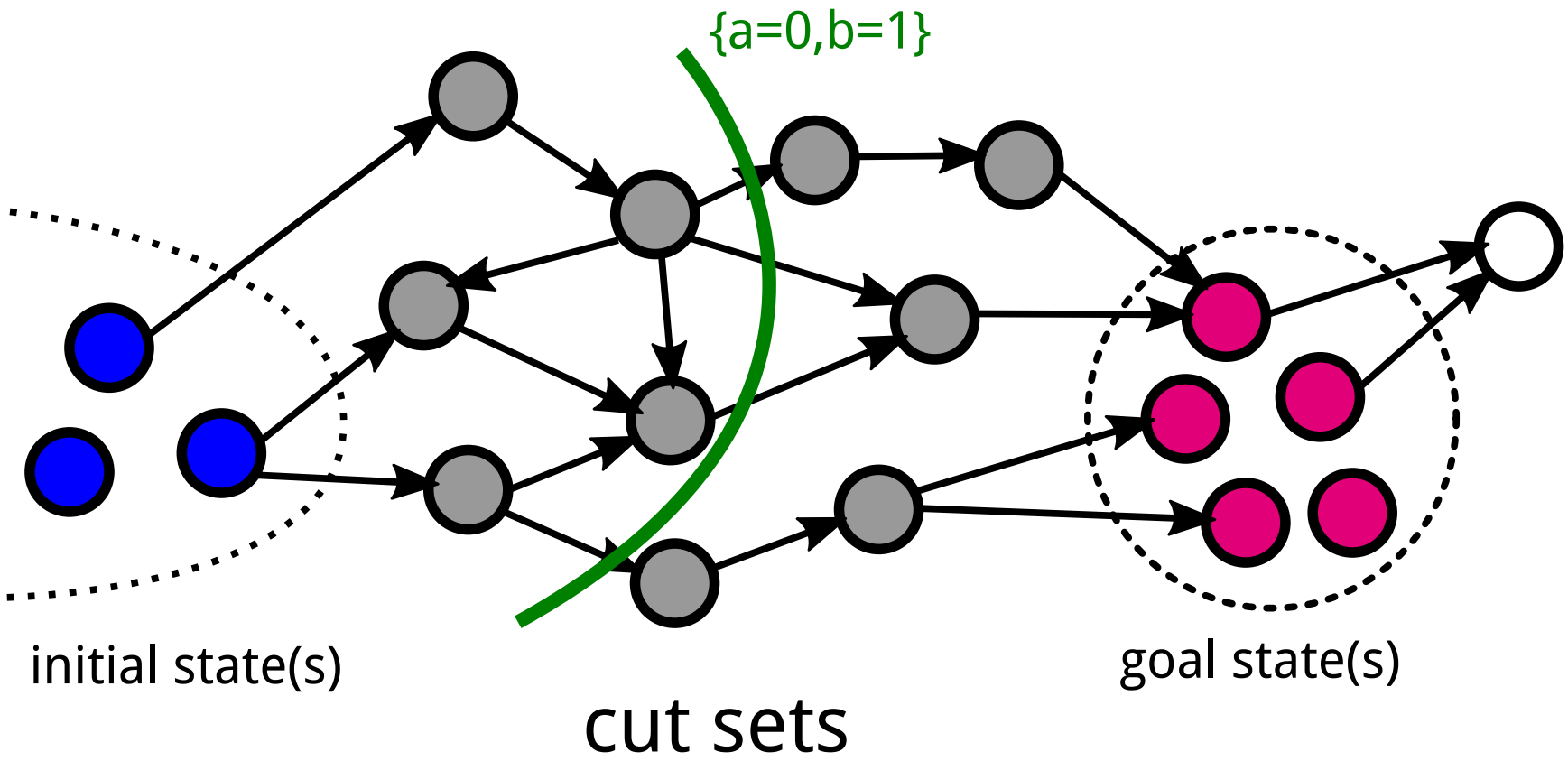



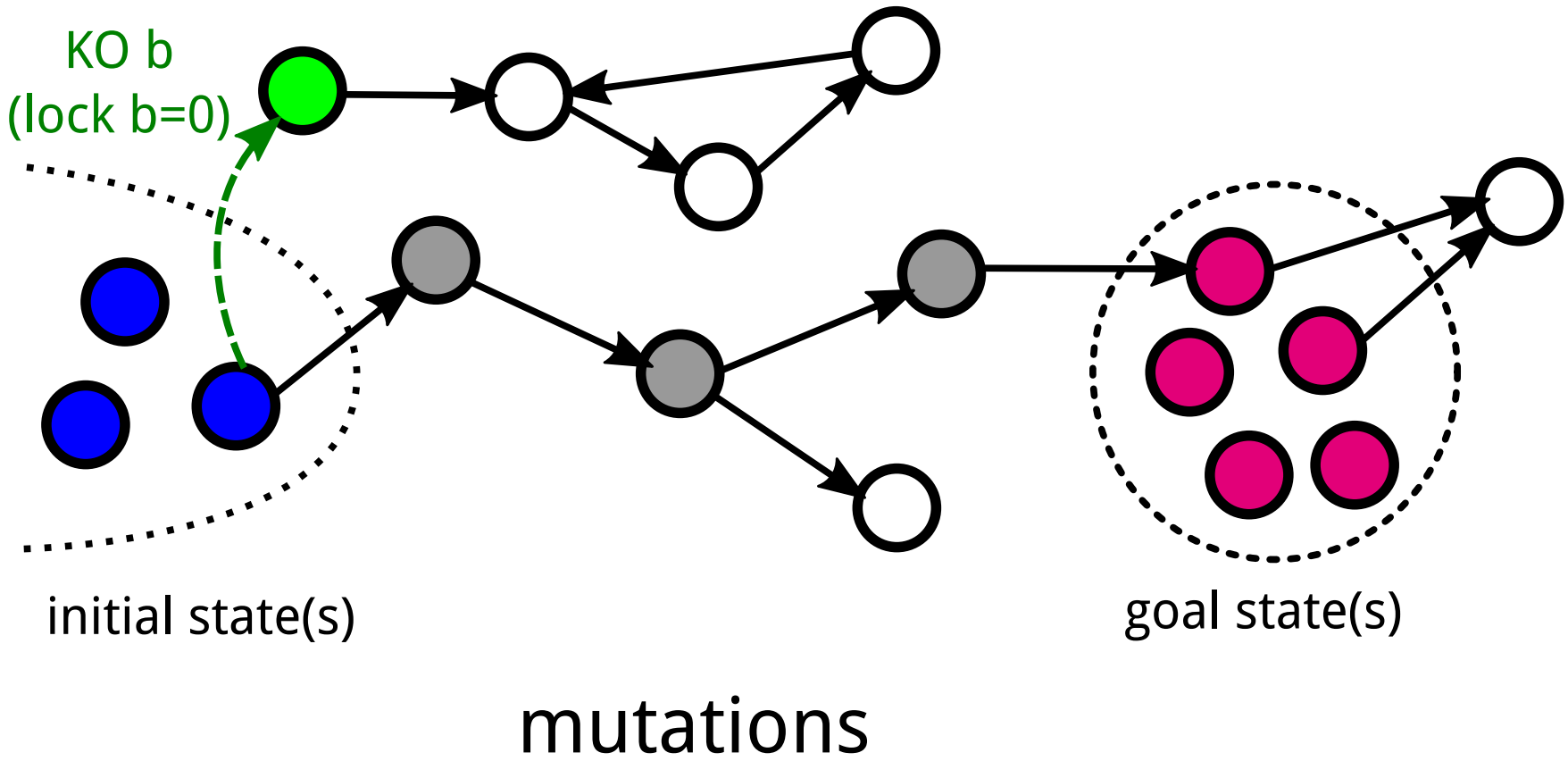




**Software Pint** - `http://loicpauleve.name/pint`
Scalability : networks with 100 - 10,000 components

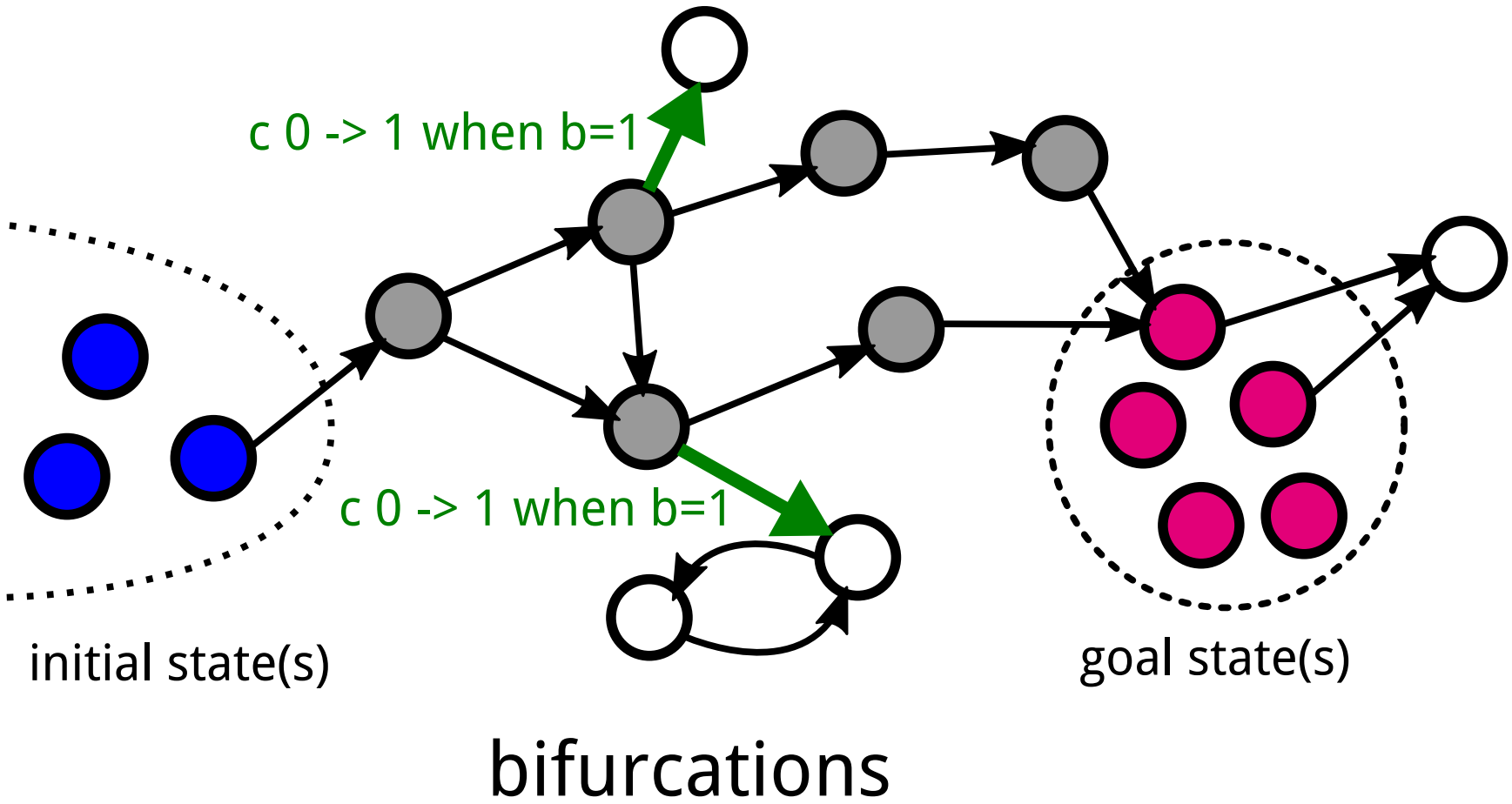
## Causal analysis for transient reachability



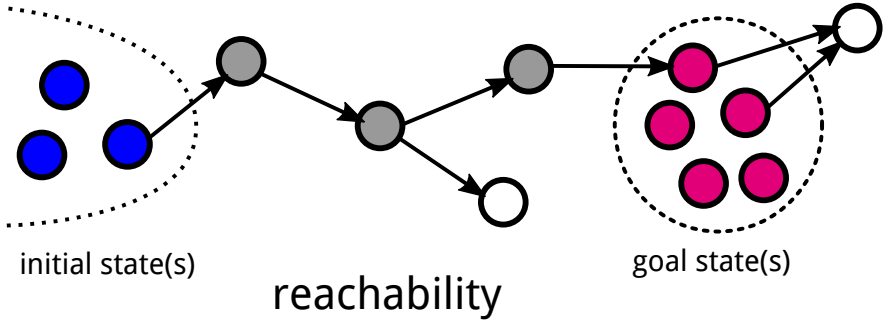
## Causal analysis for transient reachability

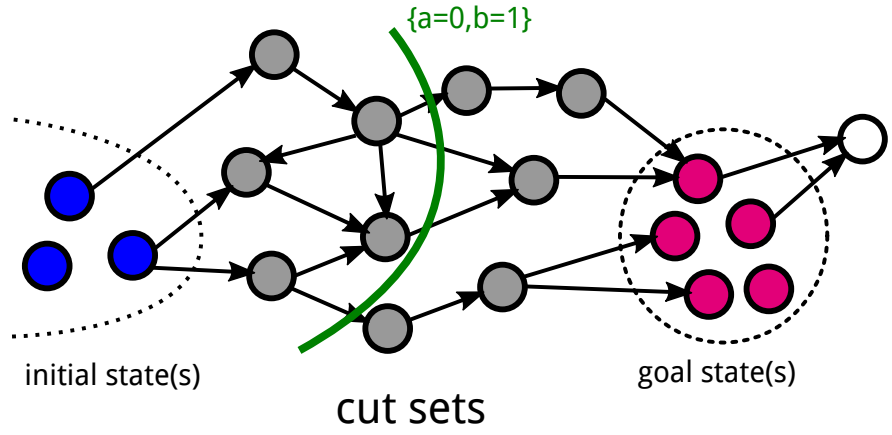


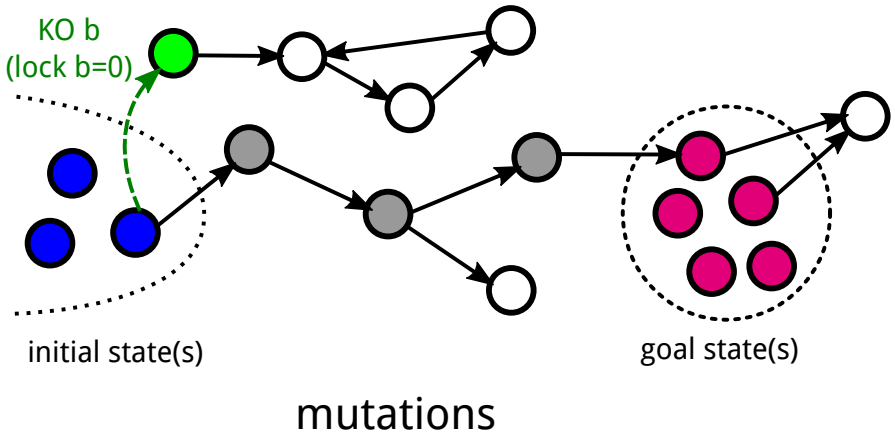
## Causal analysis for transient reachability

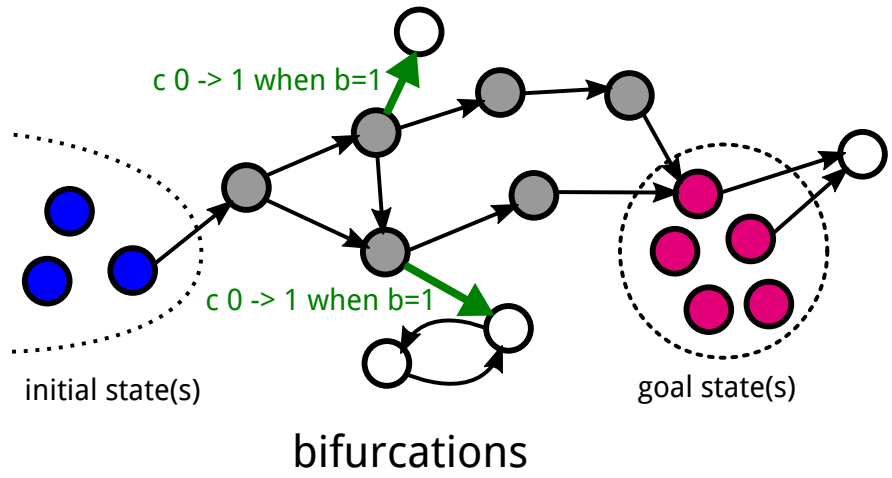


# Causal analysis for transient reachability



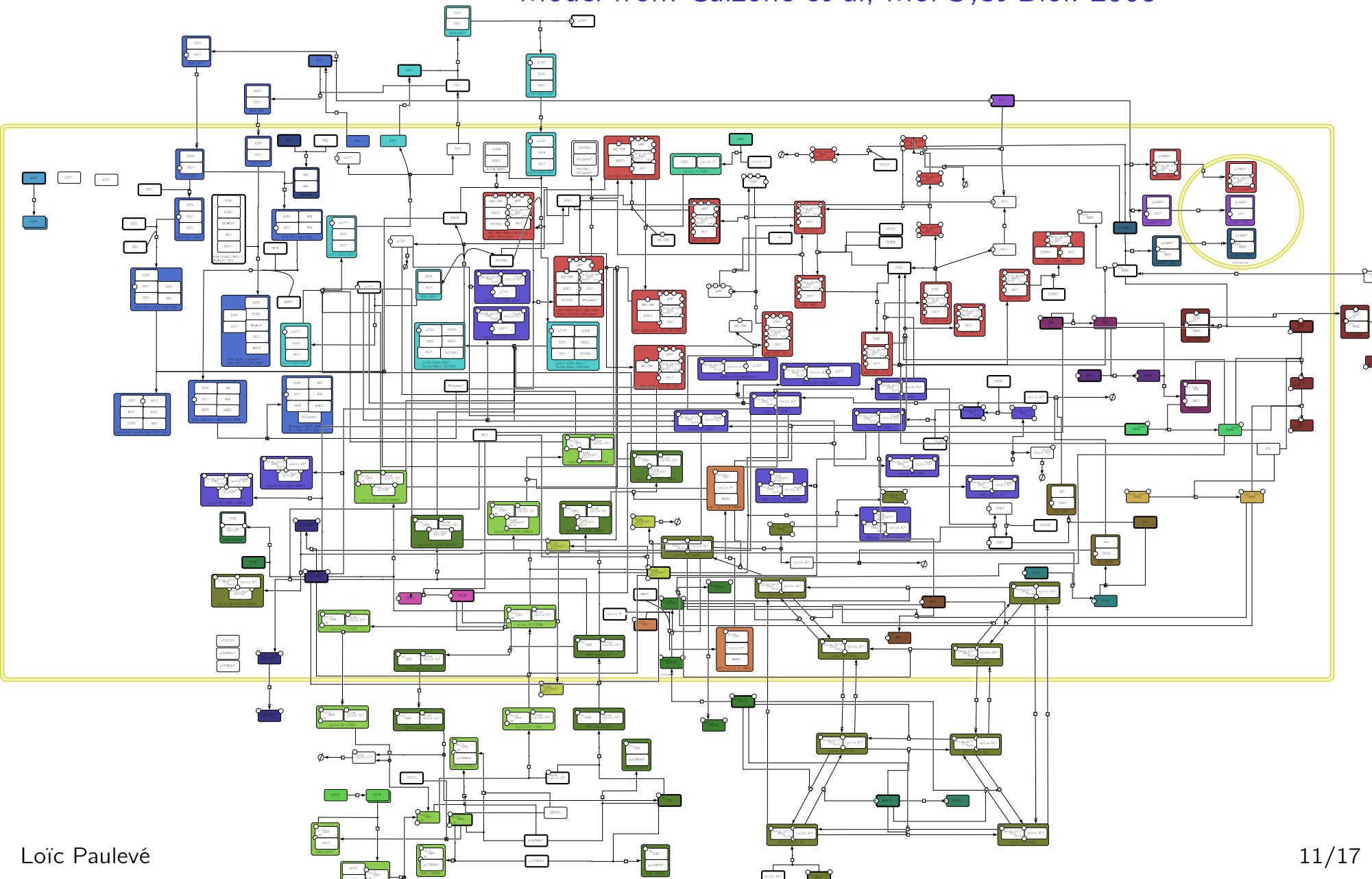




**Software Pint** - `http://loicpauleve.name/pint`
Scalability : networks with 100 - 10,000 components

# Example : Cell cycle control by RB/E2F

## Model from Calzone et al, Mol Syst Biol. 2008

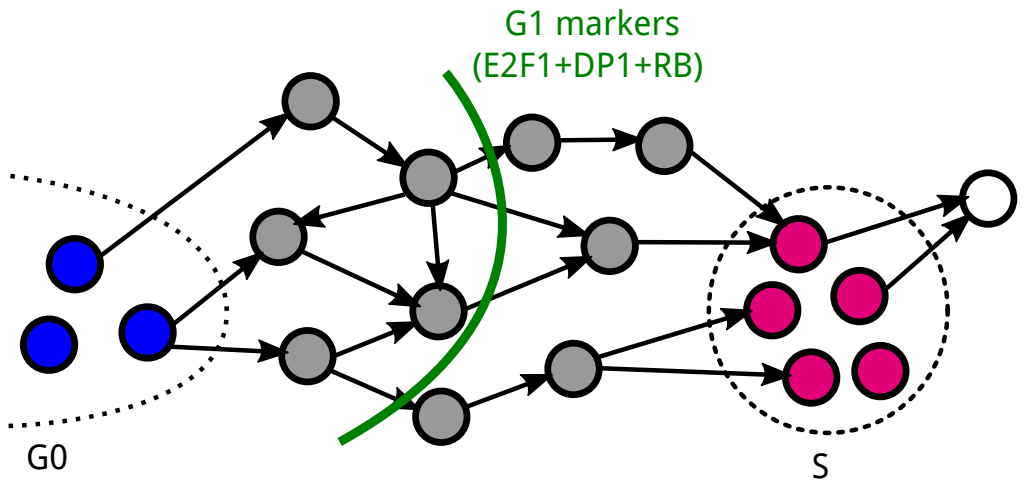


# Example : Cell cycle control by RB/E2F

Joint work with A. Rougny and C. Froidevaux

**Checking the sequence of phases**

- Are all phases reachable from G0 ?
- Are phase $n$ markers cut sets for reaching phase $n+1$ ?



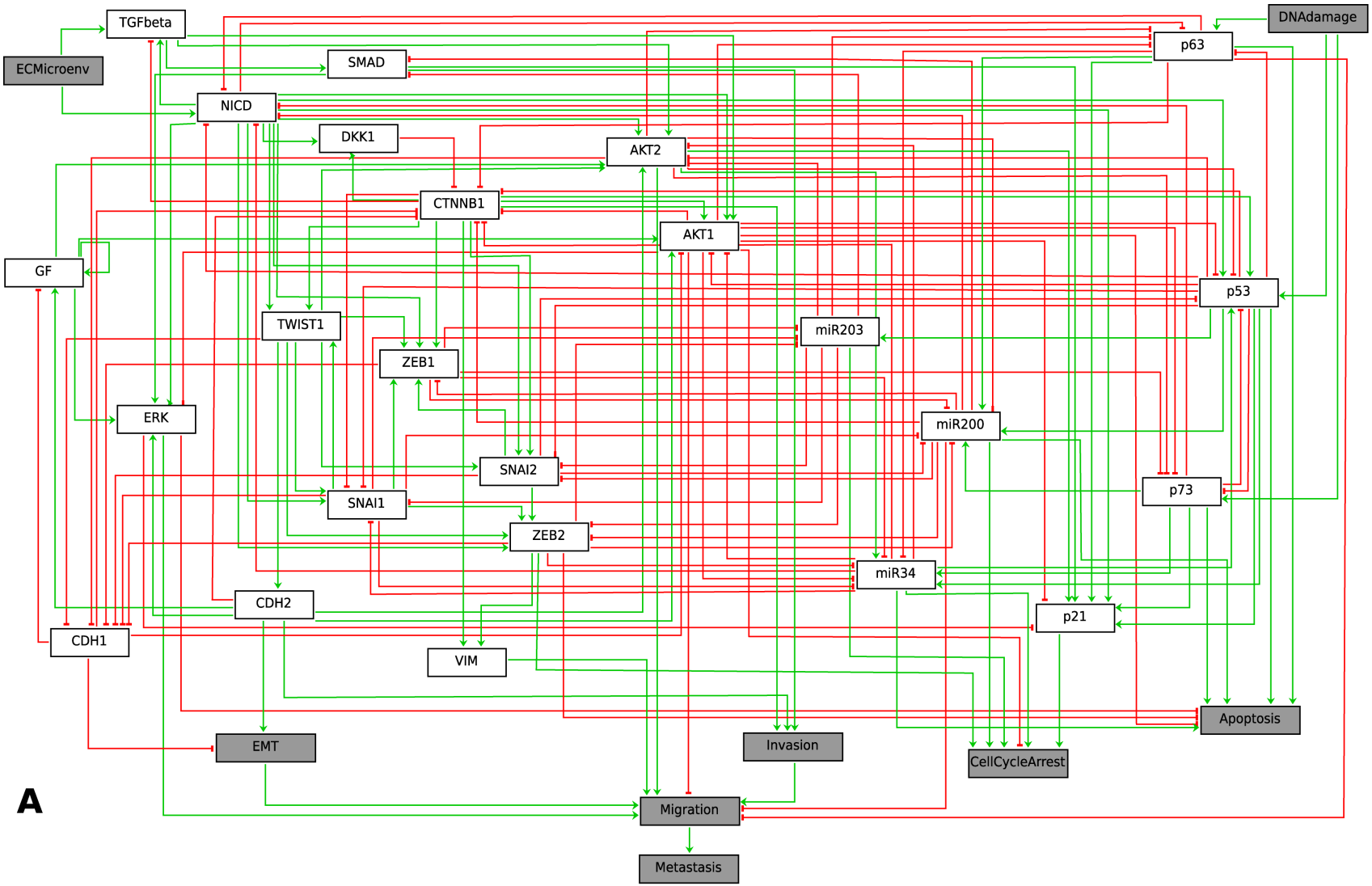

**Results of formal analysis**

Model : $\approx 300$ components, i.e., $\approx 2^{300}$ states. . . tractable only with causal analysis !

- The original map does not enforce the sequence of phases
- $\Rightarrow$ can be fixed by narrowing (known) transcriptional effects of E2F1

# Example : mutations preventing apoptosis

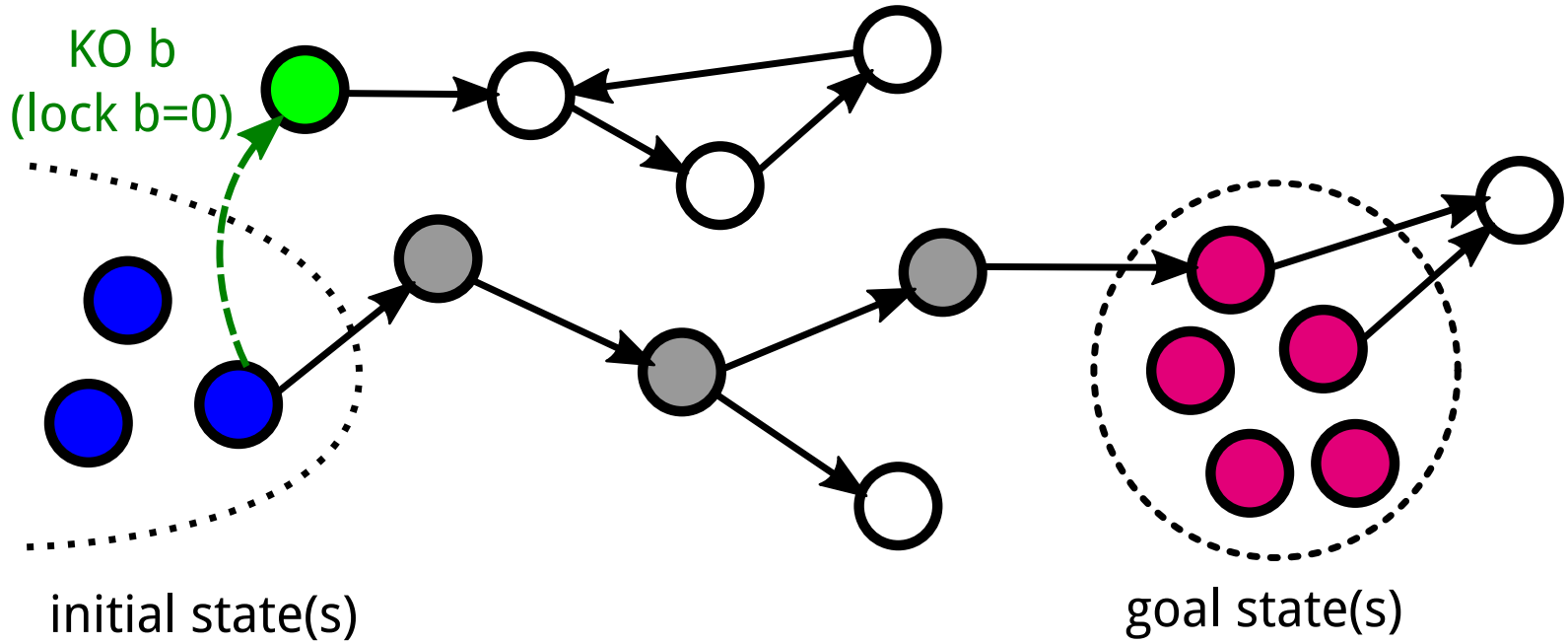
## Model from Cohen et al, Plos Comp Bio 2015



# Example : mutations preventing apoptosis

WiP w/ L. Calzone and A. Zinovyev

**Formal computation of mutations which disable apopotosis**



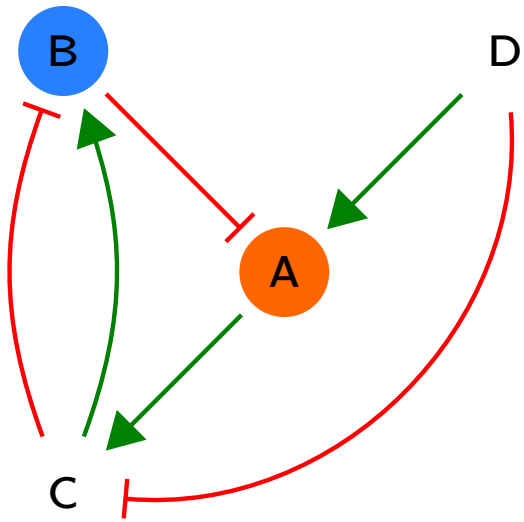

Causal analysis allows very efficient identification of mutations for reachability control

On-going work : compute temporal mutations, i.e., sequence of mutations in time.

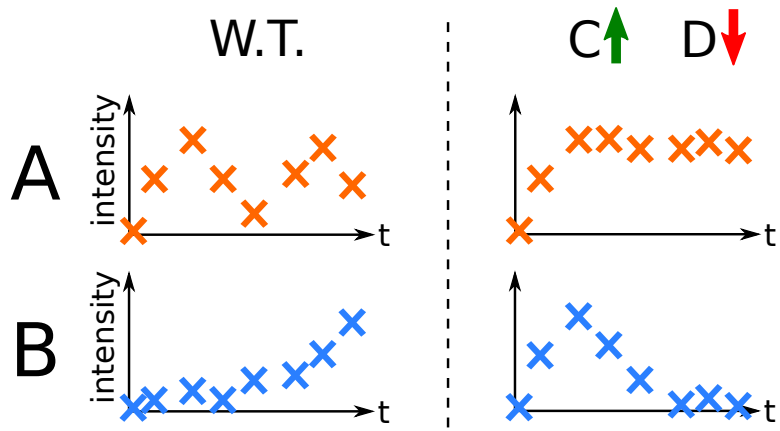
# Example : model identification

w/ Anne Siegel, Carito Guziolowski, Max Ostrowski

**Prior Knowledge Network** (PKN)
(over-approx of causal network)




**Perturbation Time Series Data**




- Identify all compatible Boolean network models
- Relies on Answer-Set Programming and approximations of reachability

Caspo-TS - `http://github.com/pauleve/caspots`

# Conclusion

**Qualitative modeling**

- Short path between knowledge and executable model
- Limit arbitrary/unobservable parameters

**Formal analysis** for models

- Validate/refute a model
- Applications : model identification, control prediction
- Efficient algorithms for automatic reasonning

**Major challenge** : deal with huge number of candidate models

- network inference lead to many networks, equivalent w.r.t. available data
- most methods take as input a single model...
- how to make convincing predictions ? (model counting...)

# Acknowledgements

**LRI, Orsay**

- Christine Froidevaux
- Adrien Rougny

**LSV/Inria Saclay, ENS Cachan**

- Stefan Haar
- Thomas Chatain
- Stefan Schwoon
- Hugues Mandon
- Juraj Kolčák

**Institut Curie**

- Laurence Calzone
- Andrei Zinovyev

**Univ. Luxembourg**

- Thomas Sauter
- Lasse Sinkonnen
- Julia Becker
- Jun Pang

**LCSB, Luxembourg**

- Antonio del Sol
- Andras Hartmann
- Andrzej Mizera
- Sacha Zickenrott

**Masaryk University**

- David Šafránek

**IRISA, Rennes** Dyliss

- Anne Siegel

**LS2N, Nantes**

- Olivier Roux
- Morgan Magnin
- Carito Guziolowski
- Maxime Folschette
- Louis Fippo Fitime

**Funding**

cnrs ANR Luxembourg National Research Fund