Causal Analysis in Computational Models of Biological Networks Dynamics

Loïc Paulevé

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

IBV - 19 May 2017

Self-introduction

CNRS Researcher in computer science lab at Univ Paris-Sud

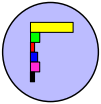
PhD from Ecole Centrale de Nantes on computational systems biology

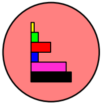
Research topic
Methods for automatic reasoning on large biological networks

From computer science to biology

- ANR HyClock (F. Delaunay): analysing detailed models of circadian clock and cell cycle.
- Starting project: ANR-FNR AlgoReCell on models and algorithms for cellular reprogramming inc. wet lab experiments (partners: Curie; Inria; Univ of Luxembourg).

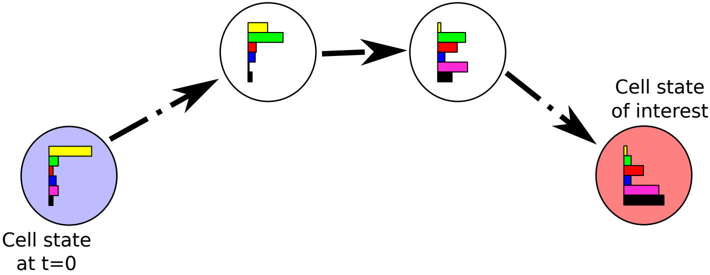
Cellular Dynamics





Initial state(s)/Goal state(s)

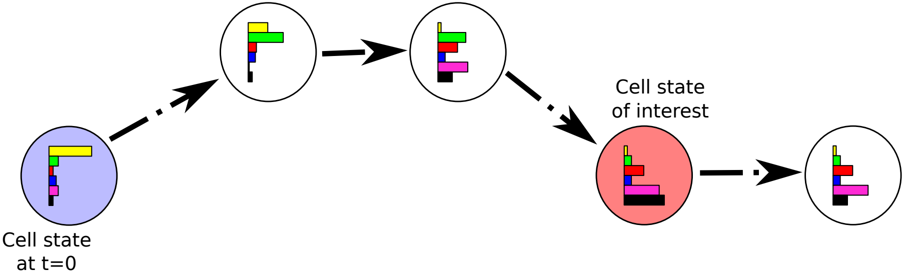
Cellular Dynamics



Initial state(s)/Goal state(s)

- Trajectory existence (reachability)

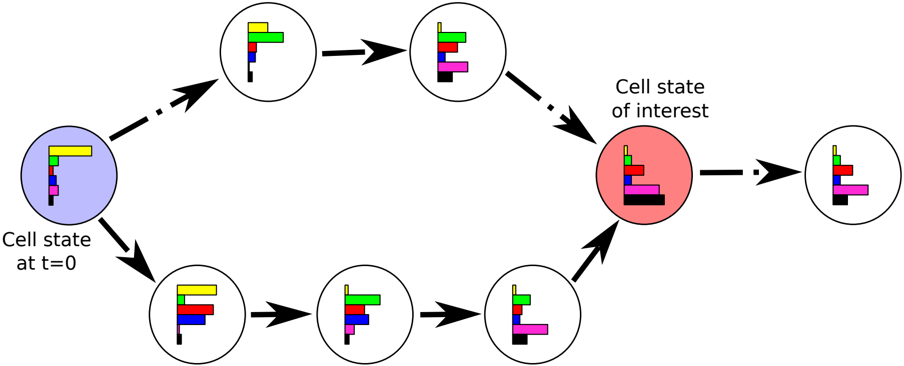
Cellular Dynamics



Initial state(s)/Goal state(s)

- Trajectory existence (reachability)

Cellular Dynamics



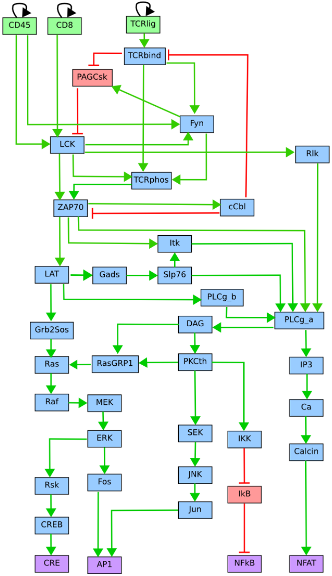
Initial state(s)/Goal state(s)

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

Outline

Outline

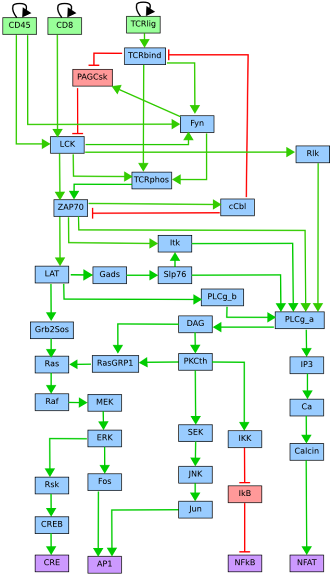
Computational models of biological networks



[Naldi et al, PLOS Comput Biol 2010]

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

Computational models of biological networks



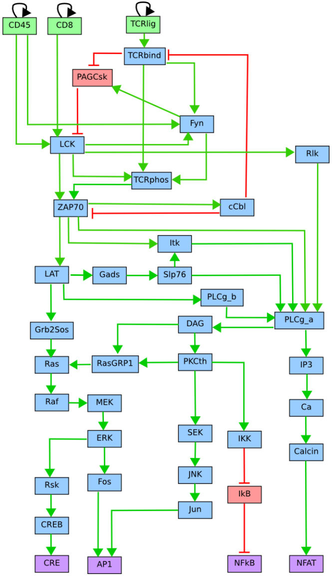
[Naldi et al, PLOS Comput Biol 2010]

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]

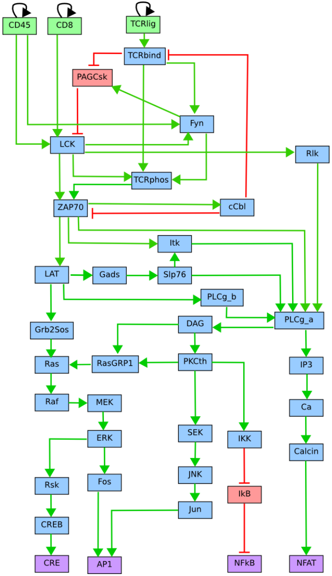
Network: account for **indirect influences** between **entities** of a system

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)

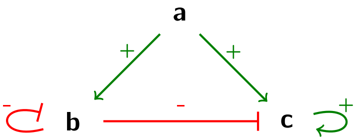
⇒ uncertainties / hypotheses remain
⇒ set of candidate models

Need for efficient methods to

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

Computational models of biological networks

Knowledge



+ Semantics

Ordinary differential equations

$$\frac{da}{dt} = -k_{da}a$$
$$\frac{db}{dt} = \frac{k_{ab}a}{1+k_{ab}a}\frac{1}{1+k_{bb}b} - k_{db}b$$
$$\frac{dc}{dt} = (\frac{k_{ac}a}{1+k_{ac}a}\frac{k_{cc}c}{1+k_{cc}c}+)\frac{1}{1+k_{bc}b} - k_{dc}c$$

Boolean network

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

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

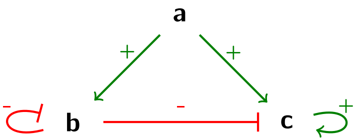
$f_c(a, b, c) =$ not b and (a or c)

Semantics

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

Computational models of biological networks

Knowledge



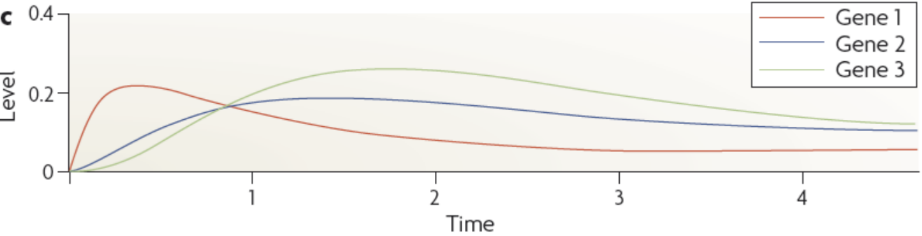
+ Semantics

Ordinary differential equations

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

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

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



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?
- 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?
- 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?
- Control prediction: perturbations which makes a behaviour impossible

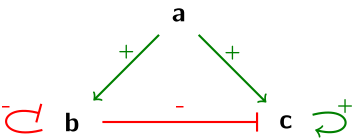
Requires a complete assessment of model capabilities:

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

Same principle to prove absence of bugs in computer programs
⇒ similar technologies, very 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)

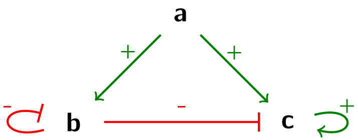
State transition graph

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

[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

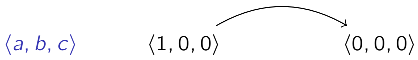


$f_a(a, b, c) = 0$

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

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

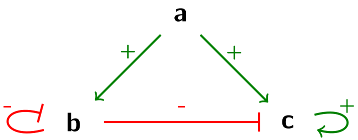
State transition graph



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

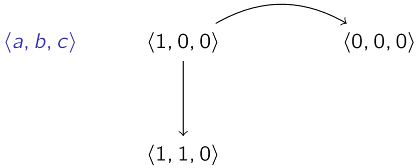
Dynamics of Qualitative Networks

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)

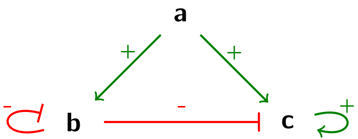
State transition graph



[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

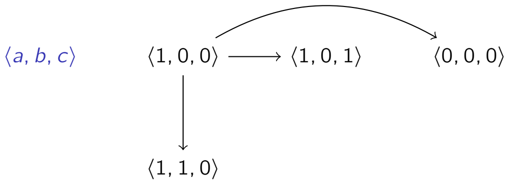


$f_a(a, b, c) = 0$

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

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

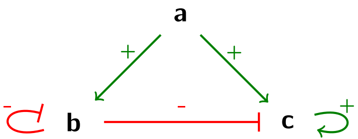
State transition graph



[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

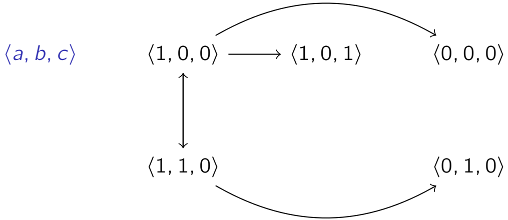


$f_a(a, b, c) = 0$

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

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

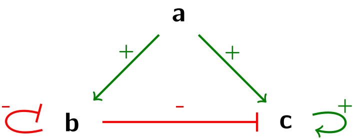
State transition graph



[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

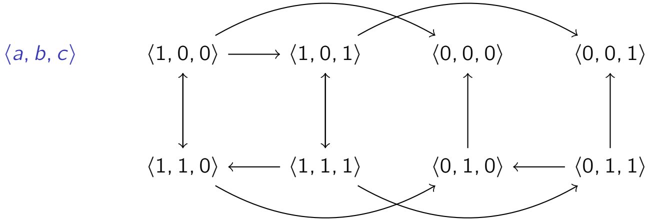


$f_a(a, b, c) = 0$

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

$f_c(a, b, c) =$ 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]

Formal Verification of Qualitative Networks

Qualitative models

- 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.

Formal Verification of Qualitative Networks

Qualitative models

- 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

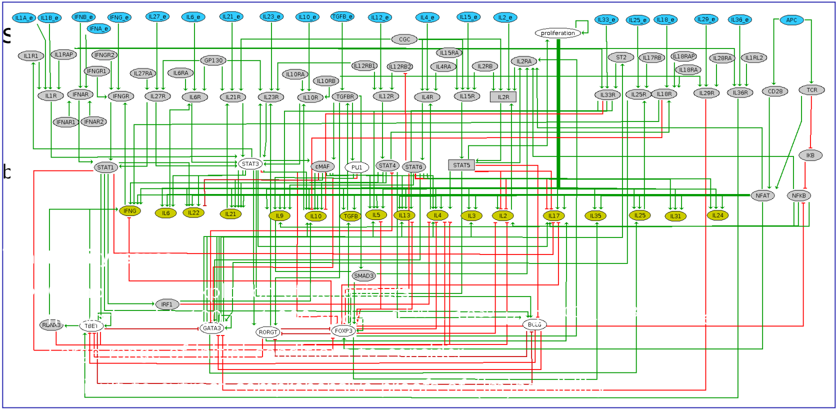
- Allows an exhaustive view of model capabilities;
- Automatic "model checking" w.r.t. specifications.

but. . .

Formal Verification of Qualitative Networks

Qualitative models

- 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.



Formal Verification of Qualitative Networks

Qualitative models

- 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

- 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.

Formal Verification of Qualitative Networks

Qualitative models

- 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

- 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.

$\Rightarrow$ **avoid building the state graph! compute something else (abstraction)**

Outline

Causal analysis in biological networks

What are the minimal causes for the changes of node states?

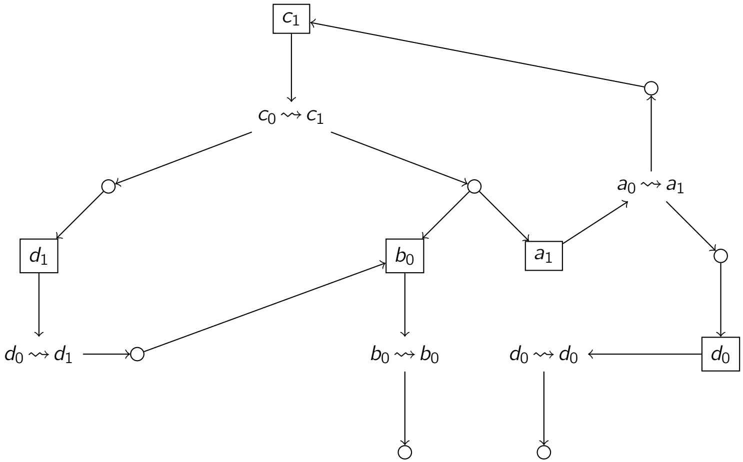
Causal analysis in biological networks

What are the minimal causes for the changes of node states?

Reason locally, i.e., only on direct regulators of the node

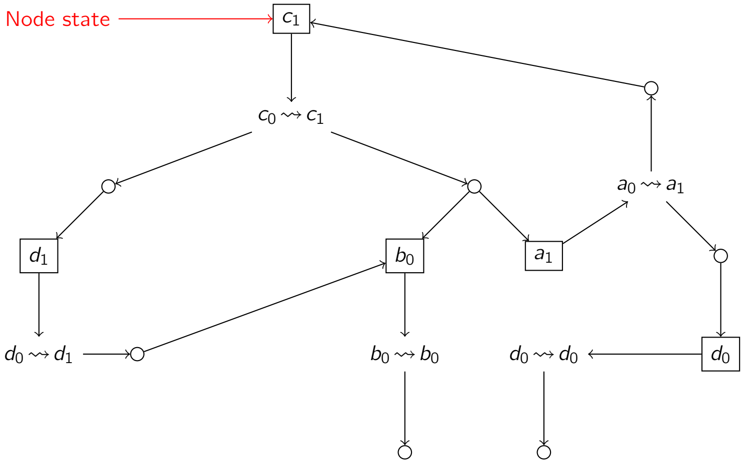
Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



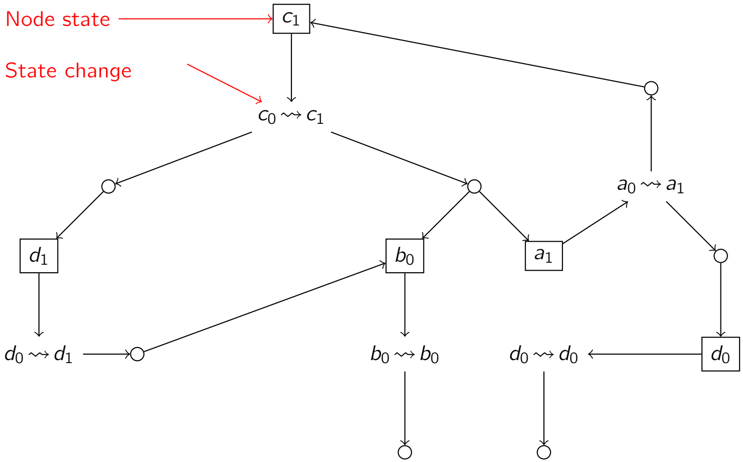
Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



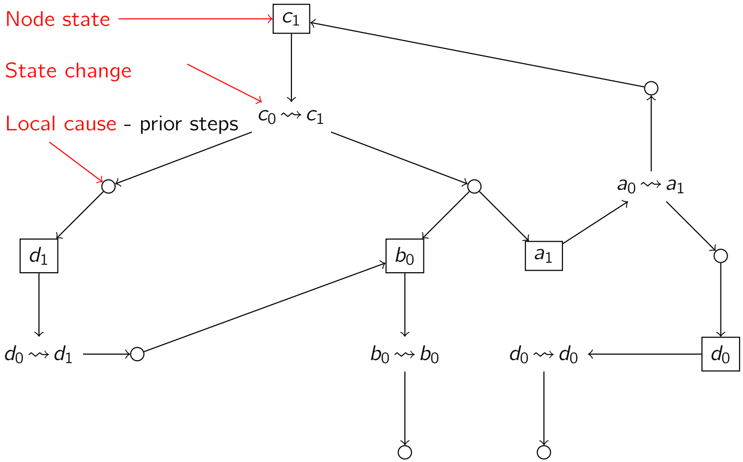
Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



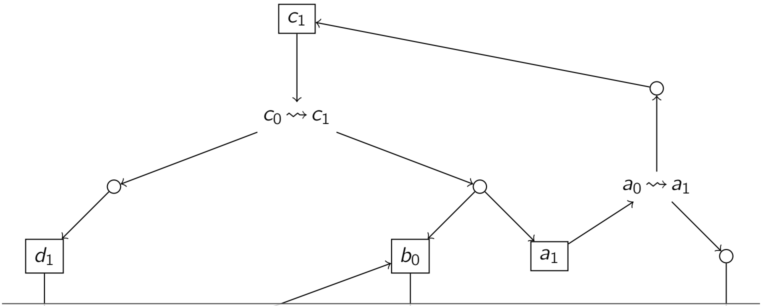
Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



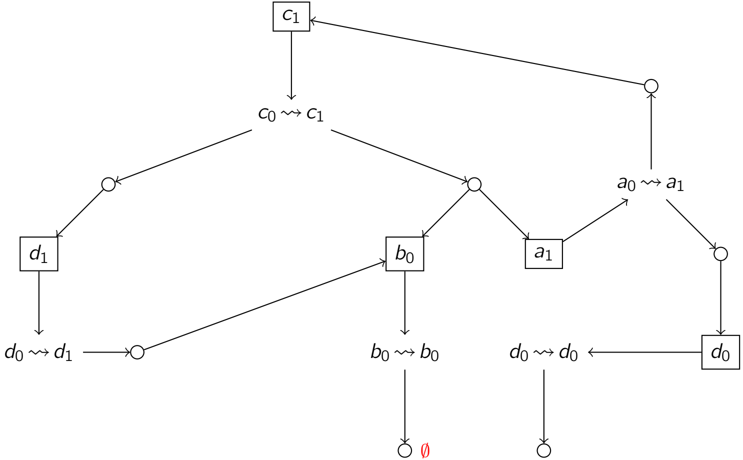
Computed automatically from formal models of networks

Size: Polynomial(nb. nodes); Exp(nb qualitative levels)
Compared to Exp(nb. nodes) for state transition graph

⇒ highly tractable for large qualitative networks

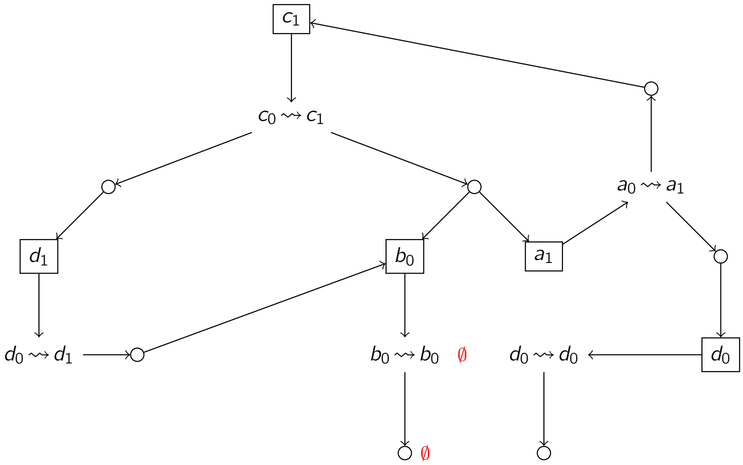
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



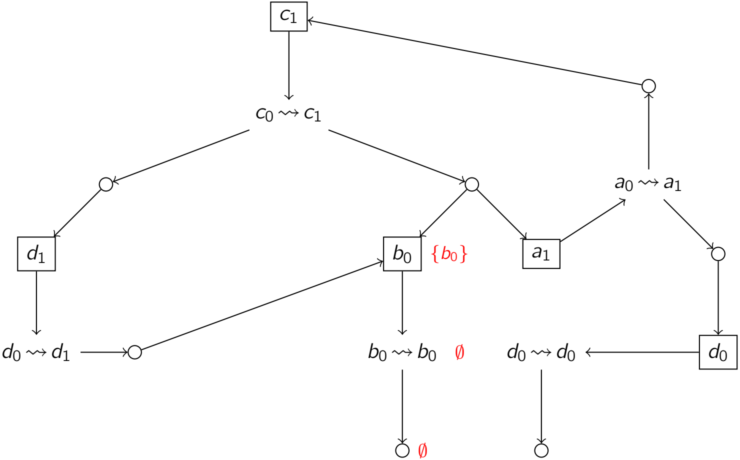
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



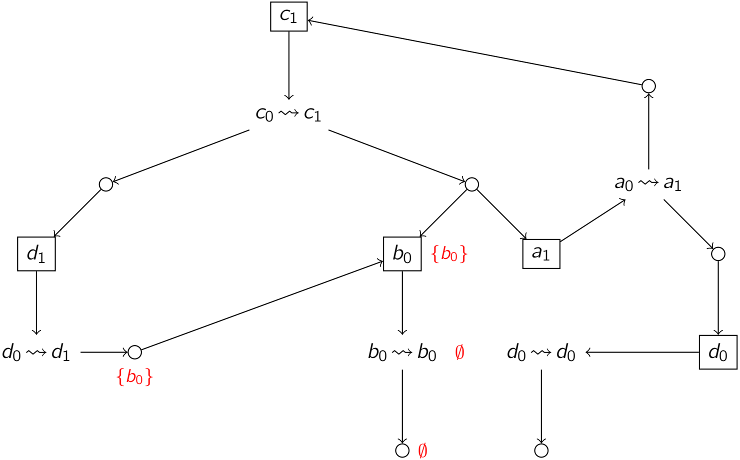
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



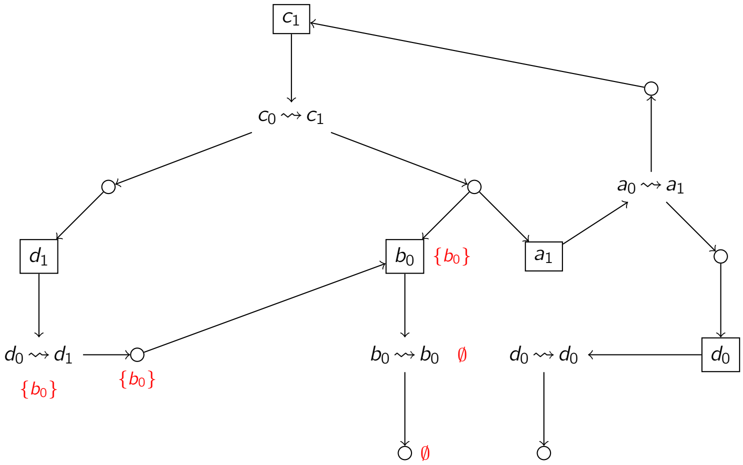
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



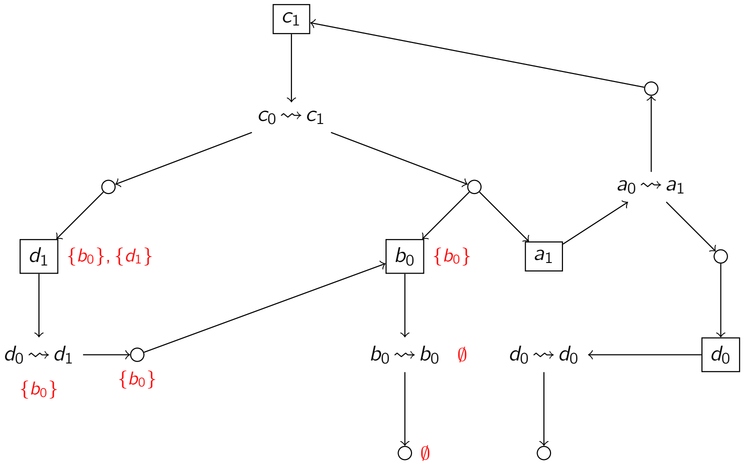
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



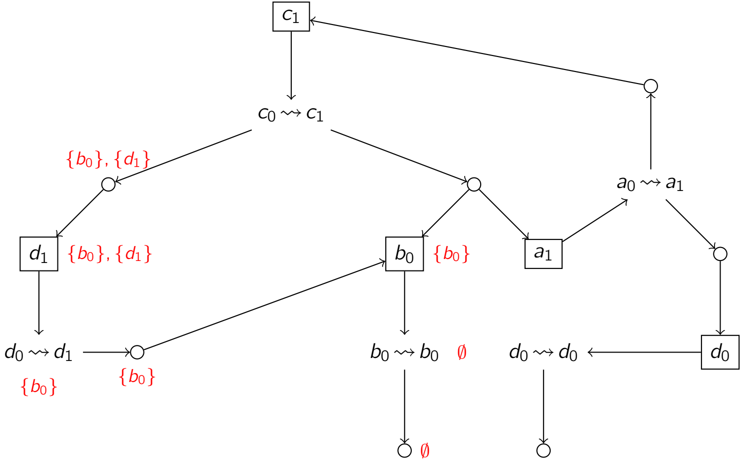
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



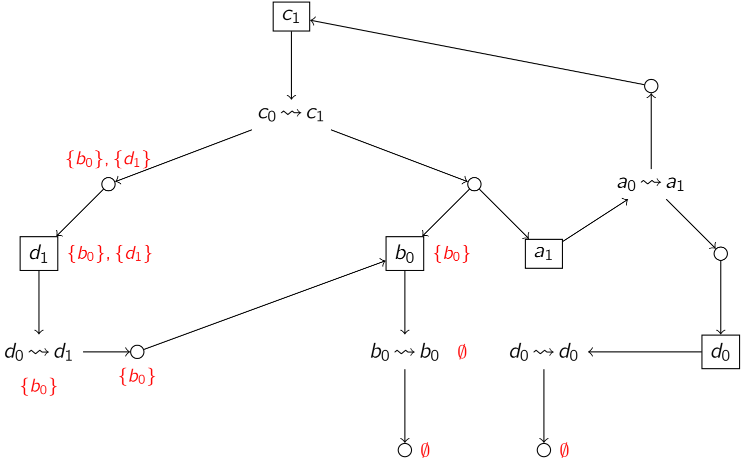
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



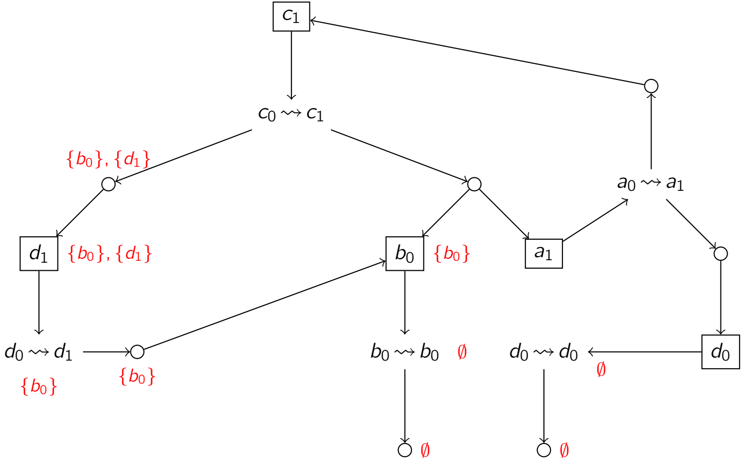
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



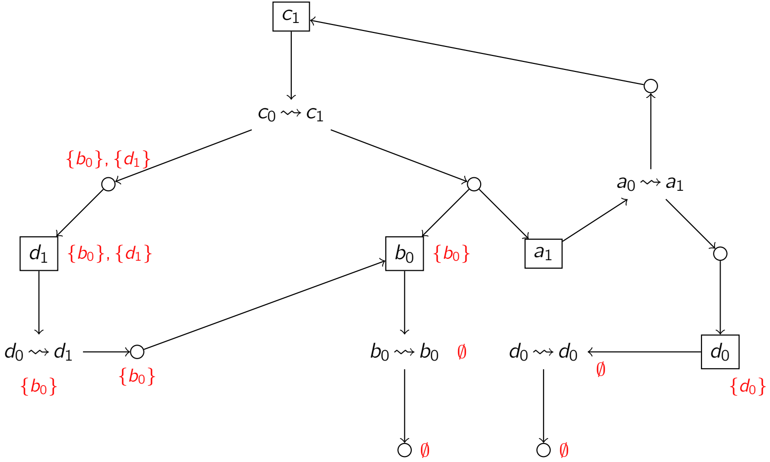
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



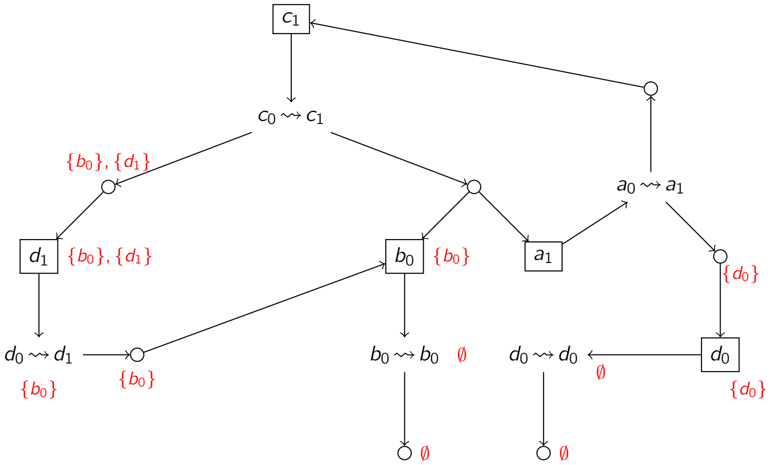
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



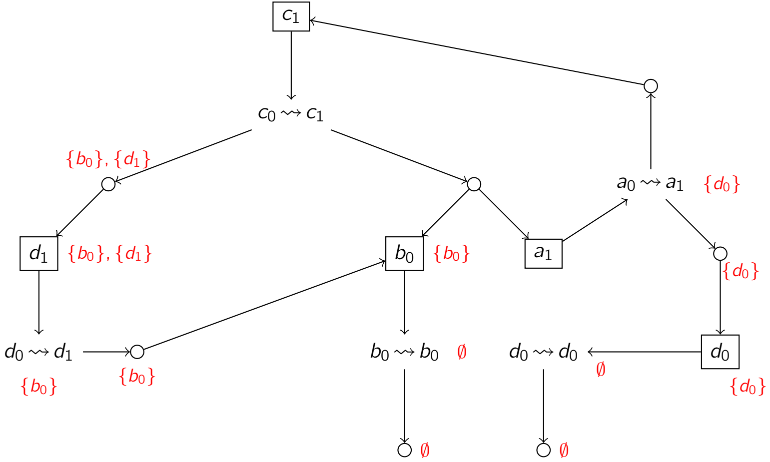
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



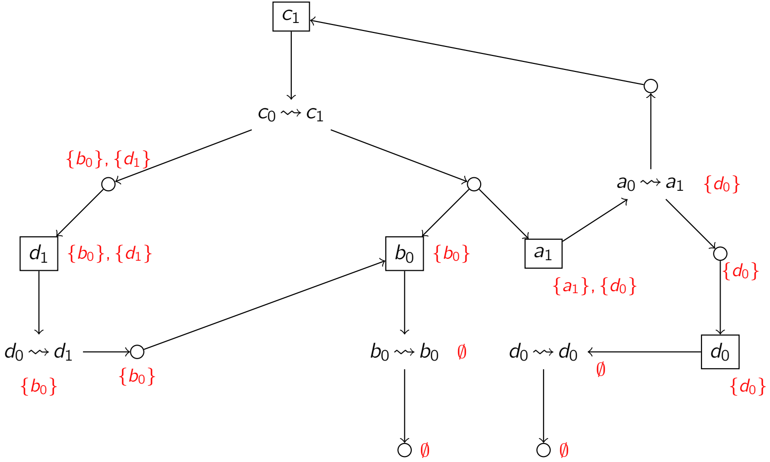
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



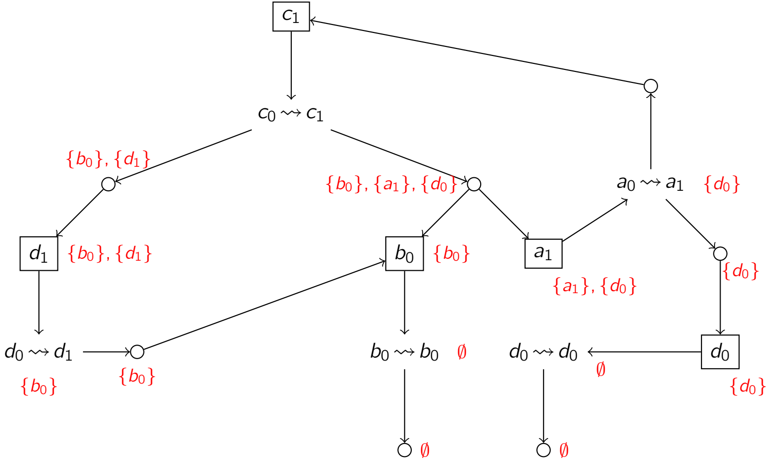
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



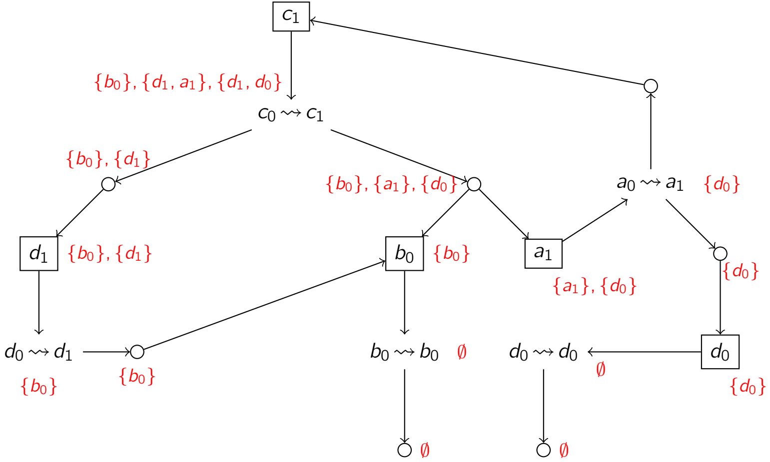
Reasoning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.



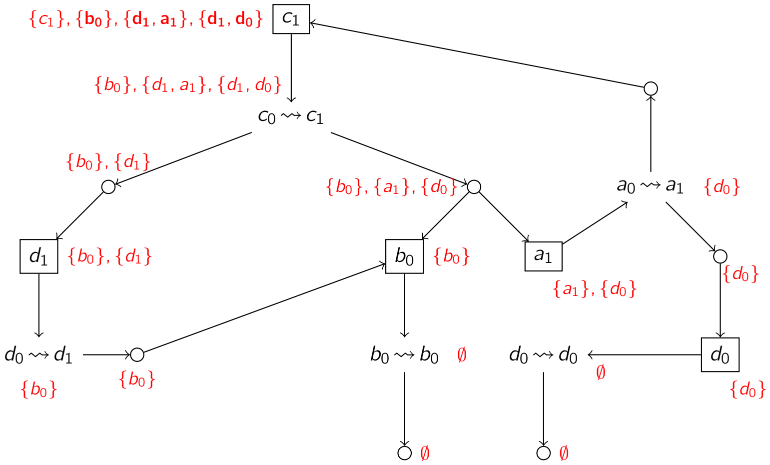
Reasonning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.

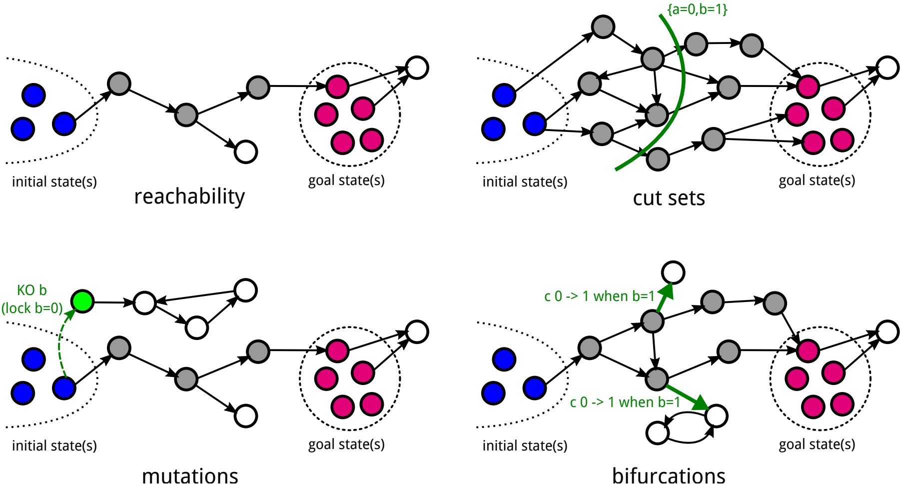


Reasoning on Local Causality Graph

- Initial state $s_0 = \{a \mapsto 0; b \mapsto 0; c \mapsto 0; d \mapsto 0\}$.

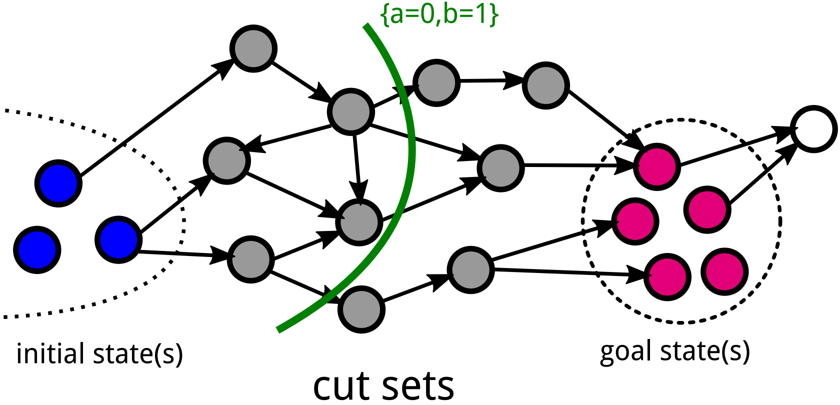


Causal analysis for transient reachability

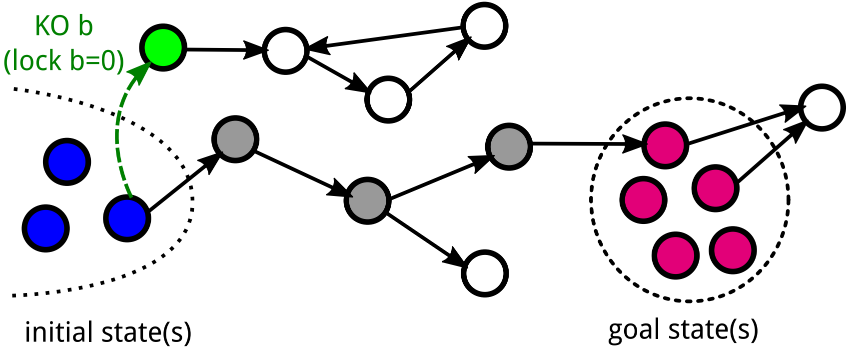


Software Pint - `http://loicpauleve.name/pint` (python interface)
Scalability: networks with 100 - 10,000 components

Causal analysis for transient reachability

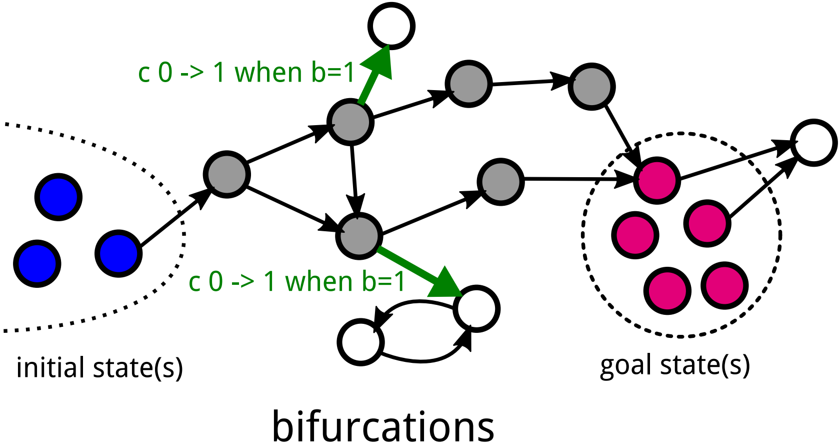


Causal analysis for transient reachability

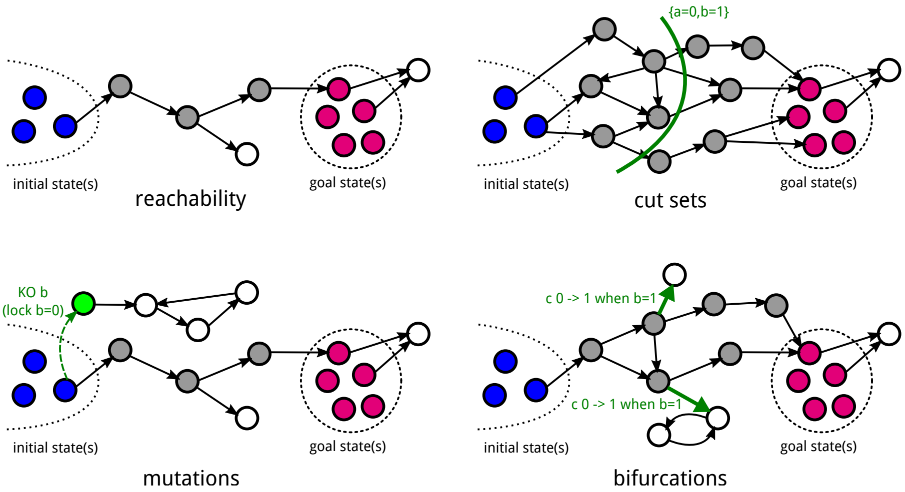


mutations

Causal analysis for transient reachability



Causal analysis for transient reachability



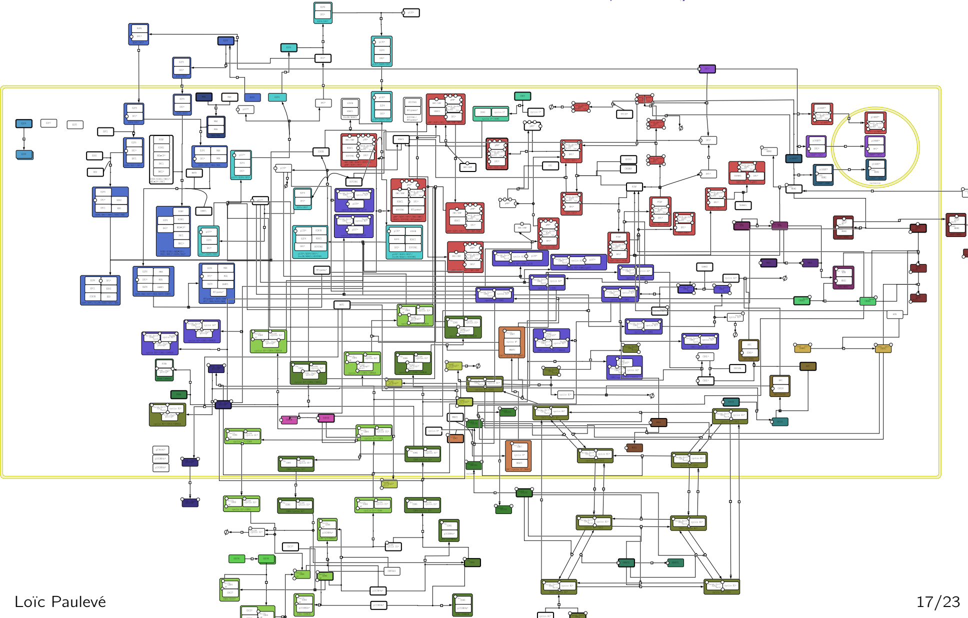
Software Pint - `http://loicpauleve.name/pint` (python interface)

Scalability: networks with 100 - 10,000 components

Outline

Example: Cell cycle control by RB/E2F

Model from Calzone et al, Mol Syst Biol. 2008

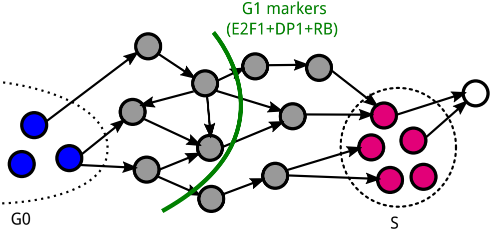


Example: Cell cycle control by RB/E2F

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 $n+1$?



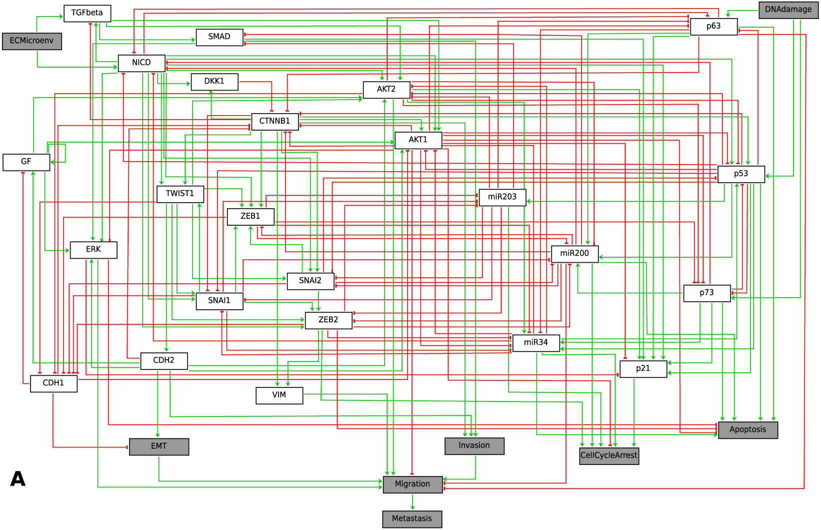
Results of formal analysis

Model: ≈ 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

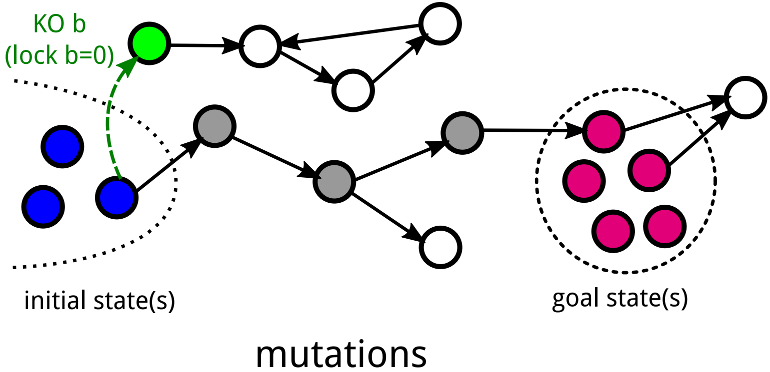
Model from Cohen et al, Plos Comp Bio 2015



Example: mutations preventing apoptosis

WiP w/ L. Calzone and A. Zinovyev

Formal computation of mutations which disable apoptosis



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: mutations preventing apoptosis

WiP w/ L. Calzone and A. Zinovyev

Formal computation of mutations which disable apopotosis

```
In [10]: wtmodel = pypint.load("http://ginsim.org/sites/default/files/SuppMat_Model_Master_Model.zginml")
         m = wtmodel.having(ECMicroenv=1,DNAdamage=0)
         m.oneshot_mutations_for_cut(Goal("Apoptosis=1")|Goal("CellCycleArrest=1"),maxsize=3,exclude={"ECMic
```

This computation is an *under-approximation*: returned mutations are all valid, but they may be non-minimal, and some solutions may be

Limiting solutions to mutations of at most 3 automata. Use `maxsize` argument to change.

```
# Running command pint-reach --json-stdout Apoptosis=1 or CellCycleArrest=1 --oneshot-mutations-for
utomata CellCycleArrest,Apoptosis,ECMicroenv -i wt.an --initial-context "ECMicroenv"=1
```

```
Out[10]: [{'AKT1': 1},
          {'CDH1': 1, 'NICD': 0},
          {'CDH2': 1, 'NICD': 0},
          {'CTNNB1': 0, 'NICD': 0},
          {'DKK1': 1, 'NICD': 0},
          {'AKT2': 1, 'ERK': 1, 'ZEB2': 0},
          {'AKT2': 1, 'ERK': 1, 'p63': 1},
          {'AKT2': 1, 'ZEB2': 0, 'p21': 0},
          {'ERK': 1, 'SNAI1': 1, 'ZEB2': 0},
          {'ERK': 1, 'SNAI2': 1, 'ZEB2': 0},
          {'ERK': 1, 'SNAI2': 1, 'p63': 1}
```

Causal analysis allows very efficient identification of mutations for reachability control

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

Outline

Conclusion

Qualitative modeling

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

Modelling causality of state changes

- Efficient algorithms for automatic reasonning
- Formal analysis of trajectories:
  - disprove a model
  - predict mutations to control the system
- Allow incomplete knowledge

Examples of directions

- Model identification: simplest models matching data and prior knowledge,
- Take into account time scales
- Algorithm for control of networks dynamics

Formal Methods for Systems Biology

Aim: **understand, analyse, control** emerging dynamics.