6e Séminaire DKM

Static Analysis with Abstract Interpretation for the Verification of Dynamical Properties of Biological Regulatory Networks

Maxime FOLSCHETTE

Dyliss team IRISA Univ Rennes, Inria, CNRS, IRISA
Dymec team IRSET Univ Rennes 1, Inserm

maxime.folschette@irisa.fr
http://maxime.folschette.name/

2018-01-18

Bio-informatics?

[De Jong, *Journal of Computational Biology*, 2002]

Bridge between **Biology** and **Computer Science**

Bio-informatics?

[De Jong, *Journal of Computational Biology*, 2002]

Bridge between **Biology** and **Computer Science**

Many fields:

- Sequencing
- Networks
- Simulation
- Experiments
- Databases
- ...

Bio-informatics?

[De Jong, *Journal of Computational Biology*, 2002]

Bridge between **Biology** and **Computer Science**

Many fields:

- Sequencing
- **Networks** ⟼ Several approaches:
  - Differential equations (quantitative)
  - Stochastic/probabilistic (quantitative)
  - Discrete graphs (qualitative)
  - Hybrid
  - ...
- Simulation
- Experiments
- Databases
- ...

Bio-informatics?

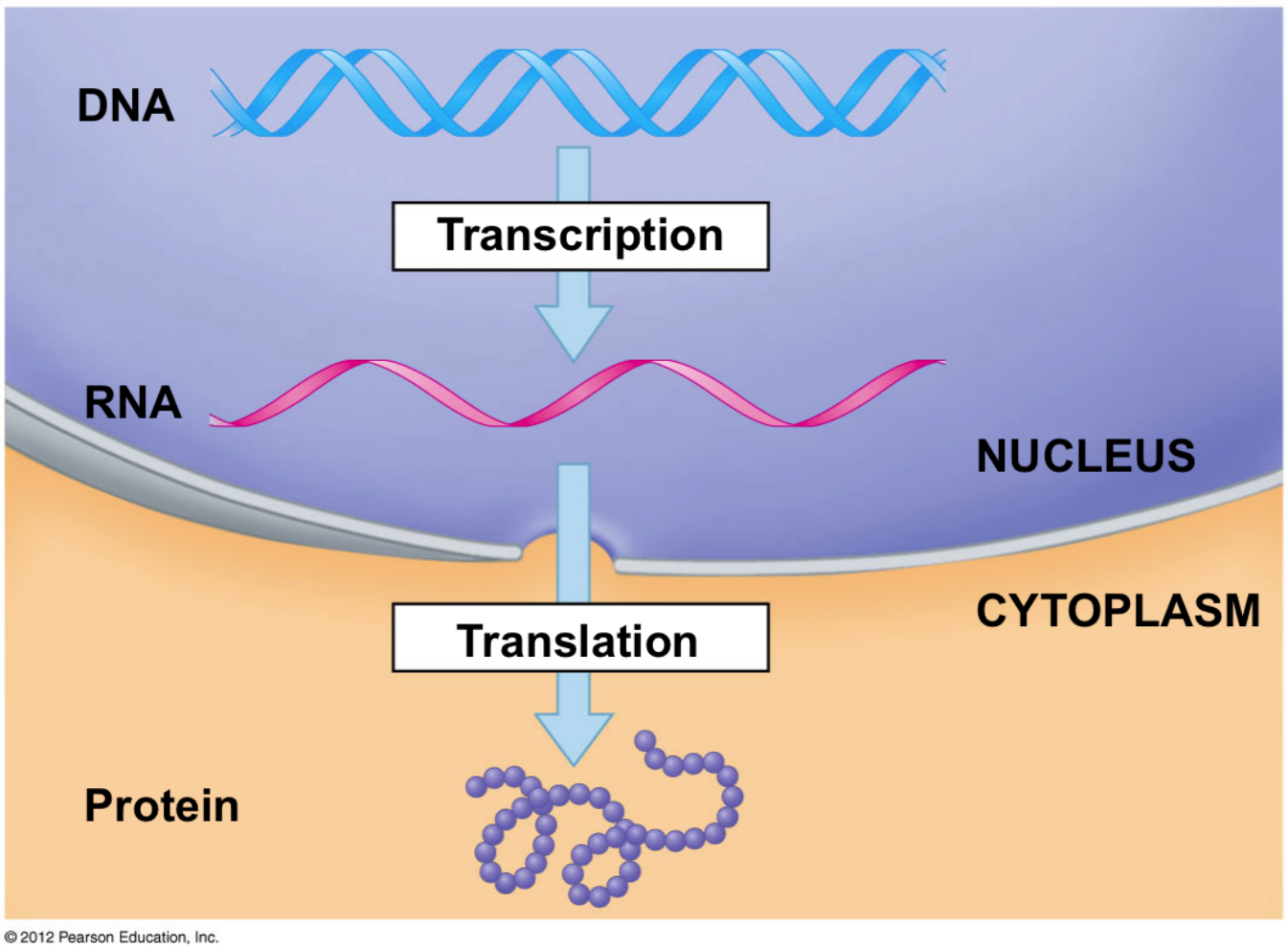
[De Jong, *Journal of Computational Biology*, 2002]

Bridge between **Biology** and **Computer Science**

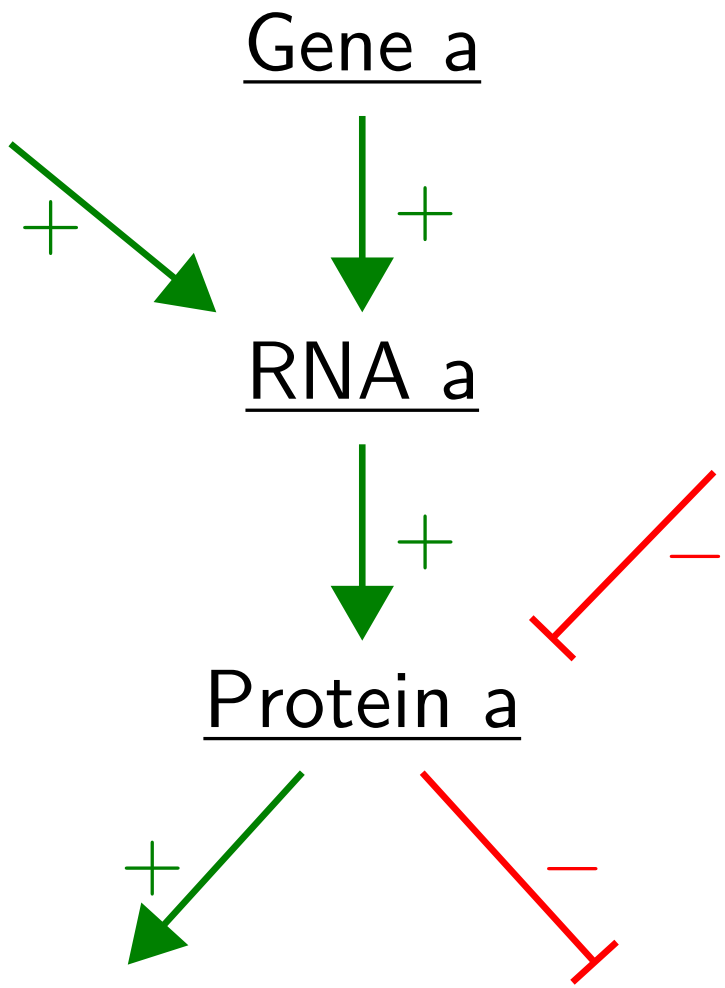
Many fields:

- Sequencing
- **Networks** ⟼ Several approaches:
  - Differential equations (quantitative)
  - Stochastic/probabilistic (quantitative)
  - **Discrete graphs (qualitative)**
  - Hybrid
  - ...
- Simulation
- Experiments
- Databases
- ...

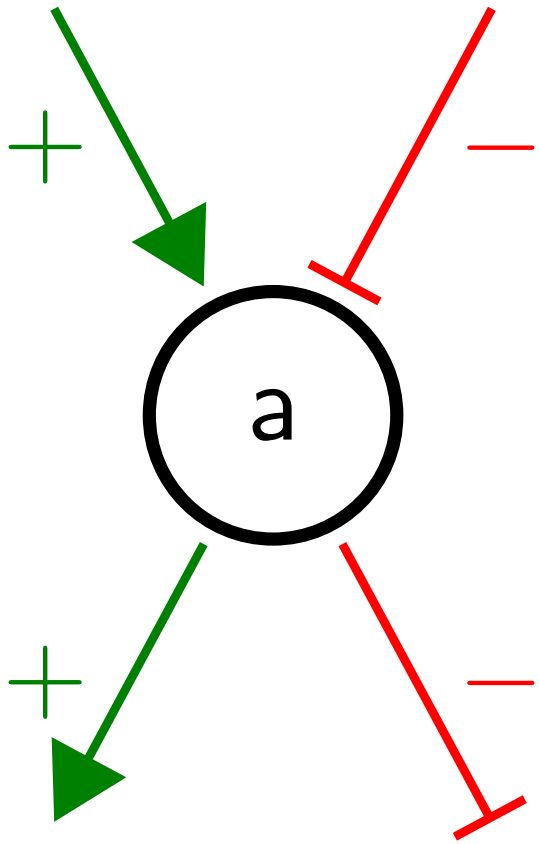
Abstraction



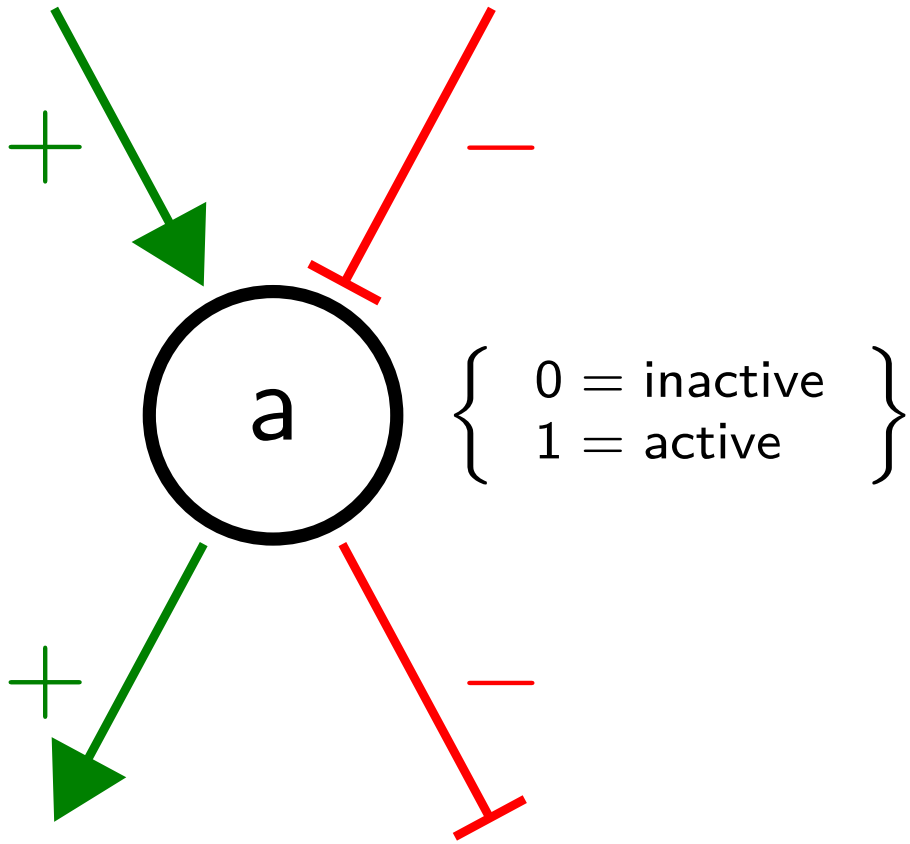
Abstraction



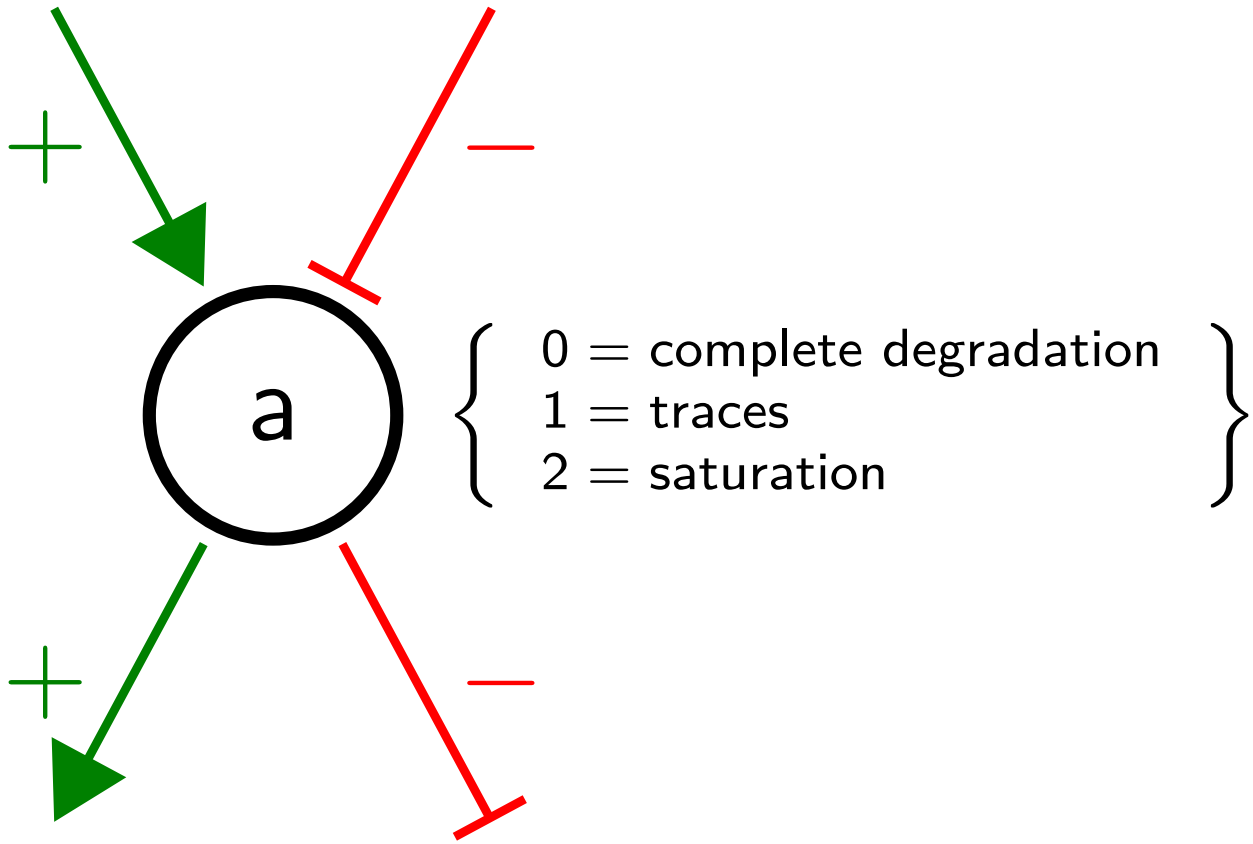
Abstraction



Abstraction

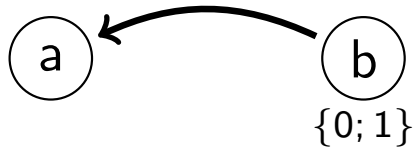


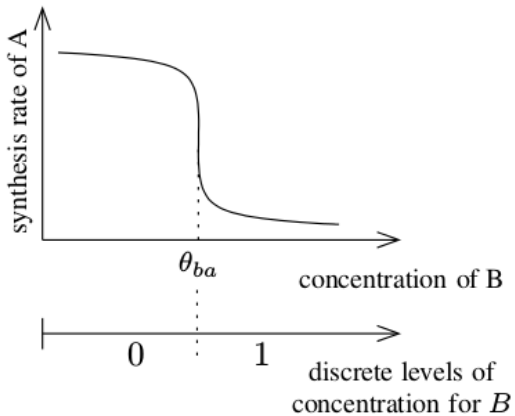
Abstraction



Discretization and Asynchronism

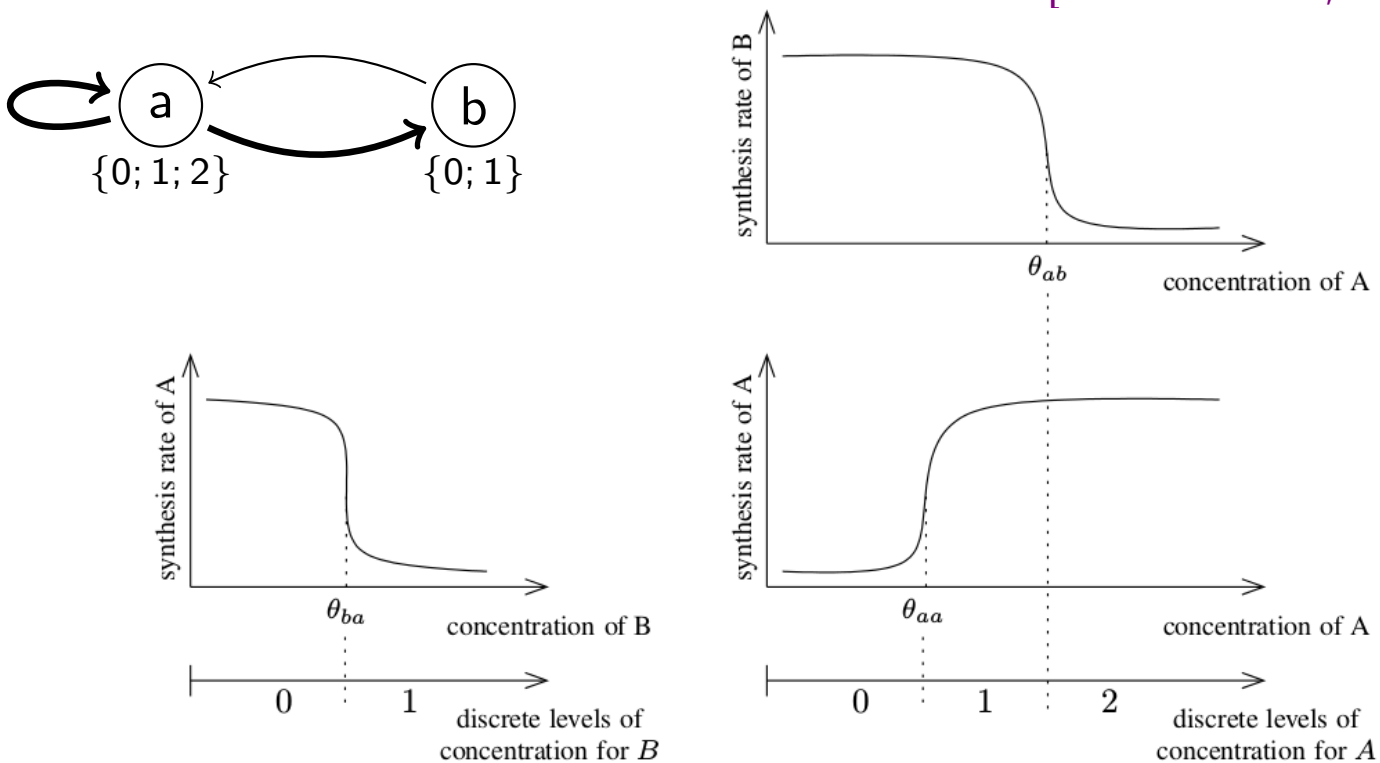
[Richard *et al.*, 2008]





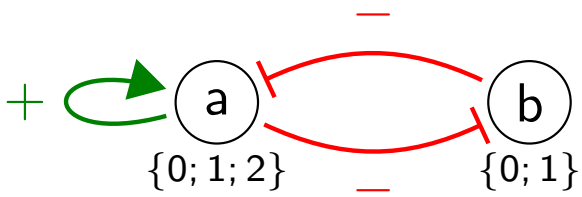
Discretization and Asynchronism

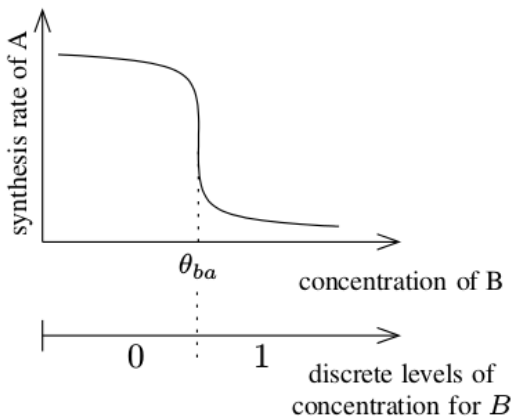
[Richard *et al.*, 2008]

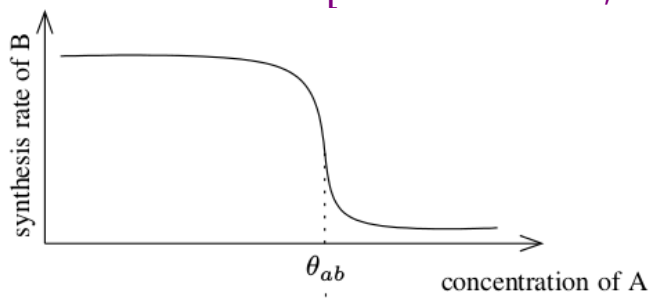


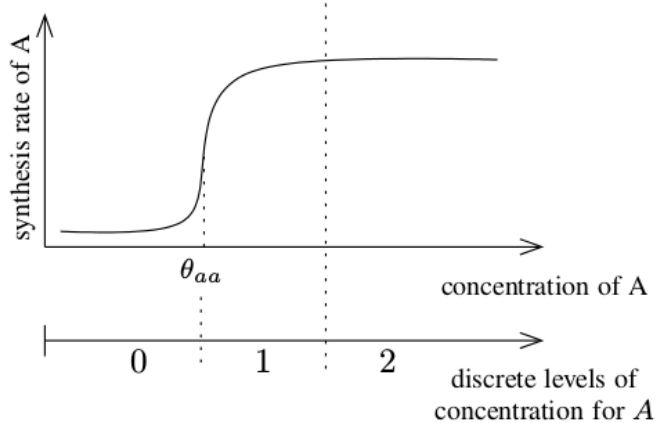
Discretization and Asynchronism

[Richard *et al.*, 2008]



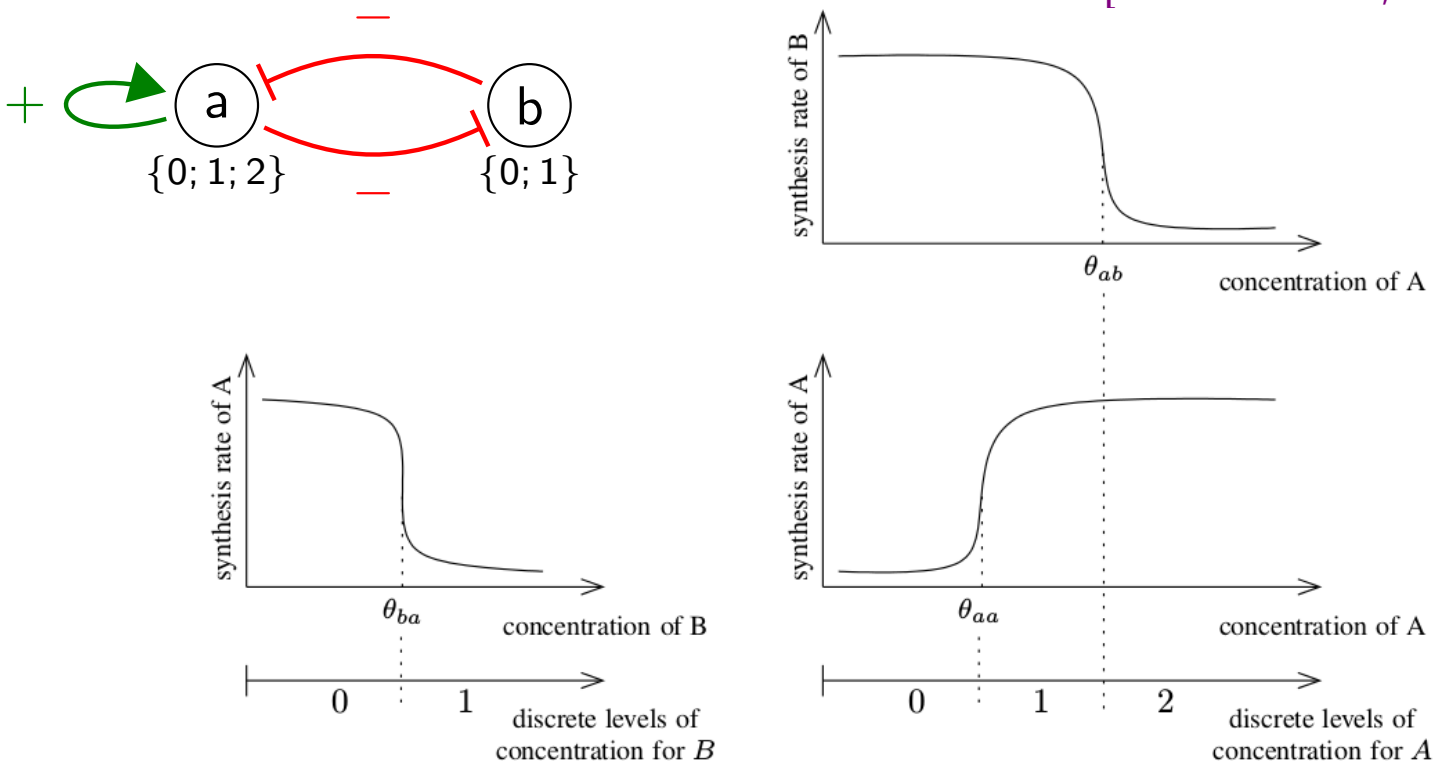






Discretization and Asynchronism

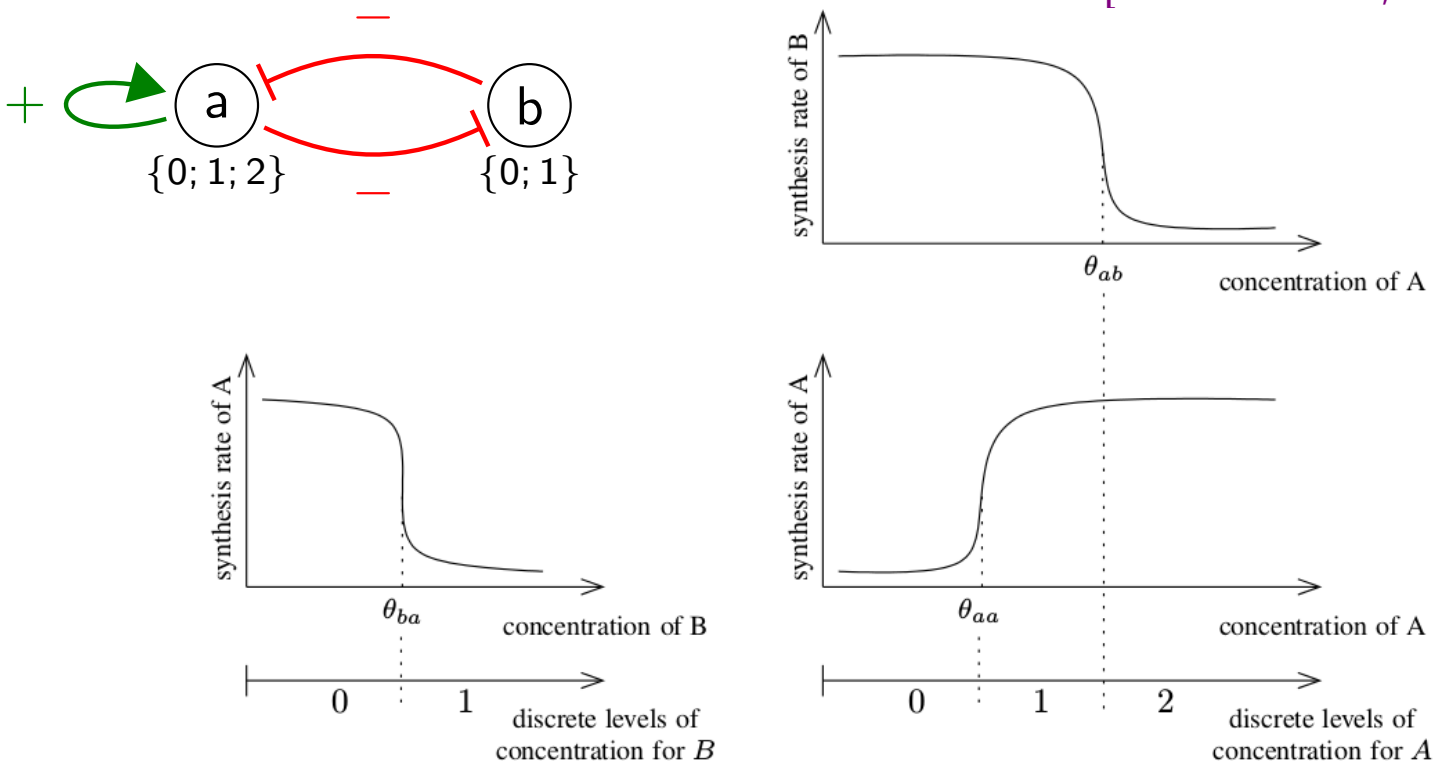
[Richard *et al.*, 2008]



- Unknown real values of concentrations or continuous activity levels
  → Abstracted as **discrete levels** and thresholds

Discretization and Asynchronism

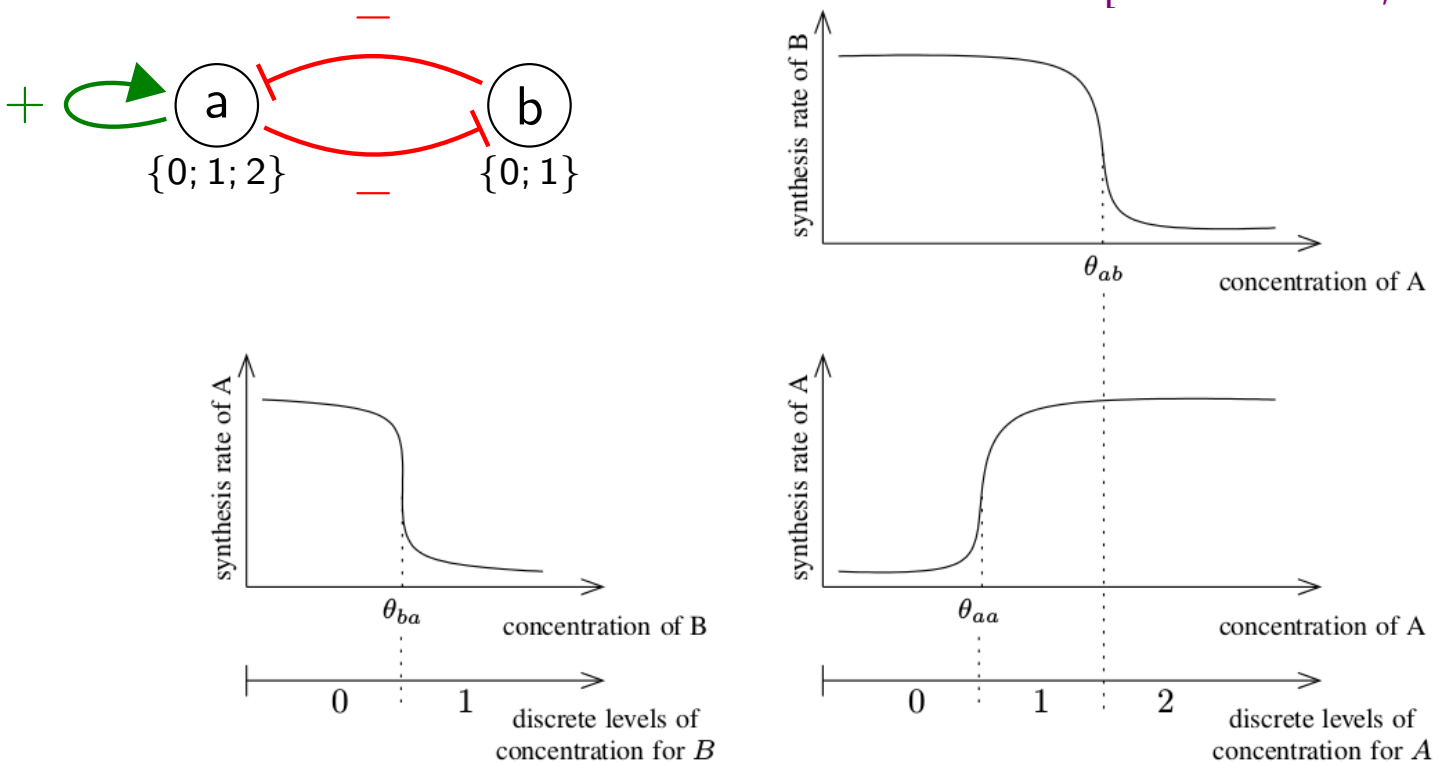
[Richard *et al.*, 2008]



- Unknown real values of concentrations or continuous activity levels
  → Abstracted as **discrete levels** and thresholds
- Continuous variations of the real values
  → **Unitary** dynamics

Discretization and Asynchronism

[Richard *et al.*, 2008]

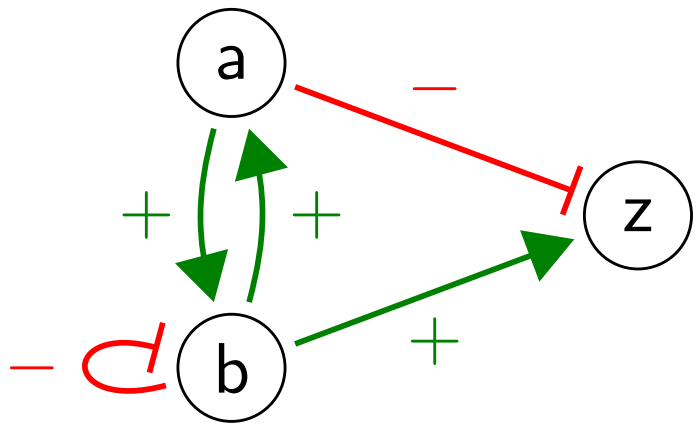


- Unknown real values of concentrations or continuous activity levels
  → Abstracted as **discrete levels** and thresholds
- Continuous variations of the real values
  → **Unitary** dynamics
- Simultaneous crossings of two thresholds never occurs
  → **Asynchronous** dynamics

Modeling of René Thomas

[Thomas, *Journal of Theoretical Biology*, 1973]

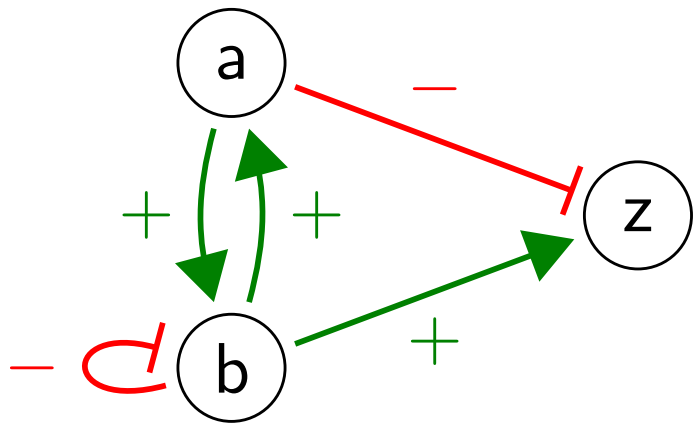
Models of interacting components (genes, proteins, ...)



Modeling of René Thomas

[Thomas, *Journal of Theoretical Biology*, 1973]

Models of interacting components (genes, proteins, ...)



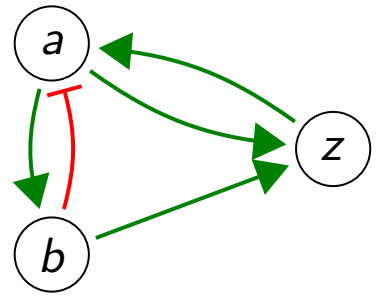
Questions:

- How does (z) **behave**?
- Is it **possible** to make (a) inactive?
- If I **knock-out** (b), what changes?

State-graph

The state-graph depicts explicitly the whole dynamics

abz			
000	010	001	011
100	110	101	111
200	210	201	211



State-graph

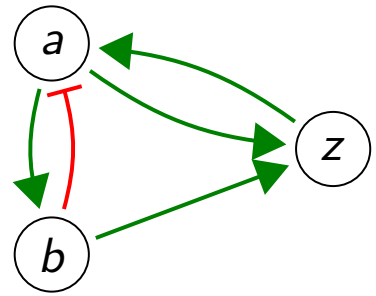
The state-graph depicts explicitly the whole dynamics

abz

000 010 001 011

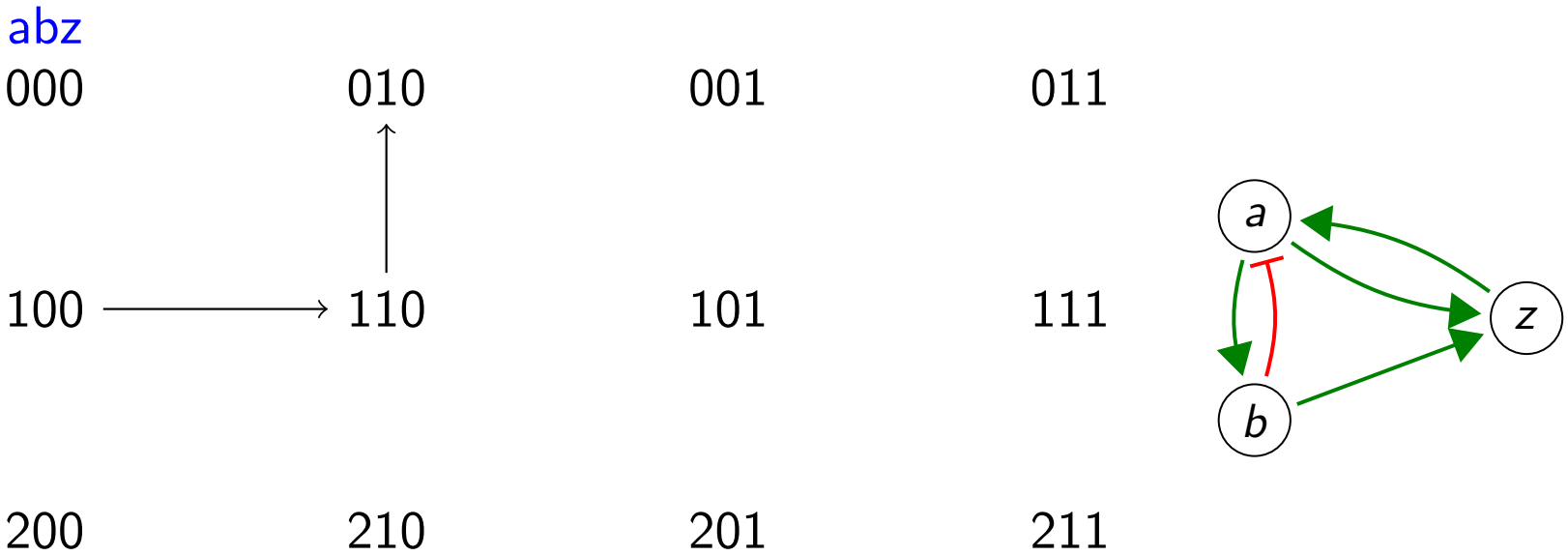
100 ⟶ 110 101 111

200 210 201 211



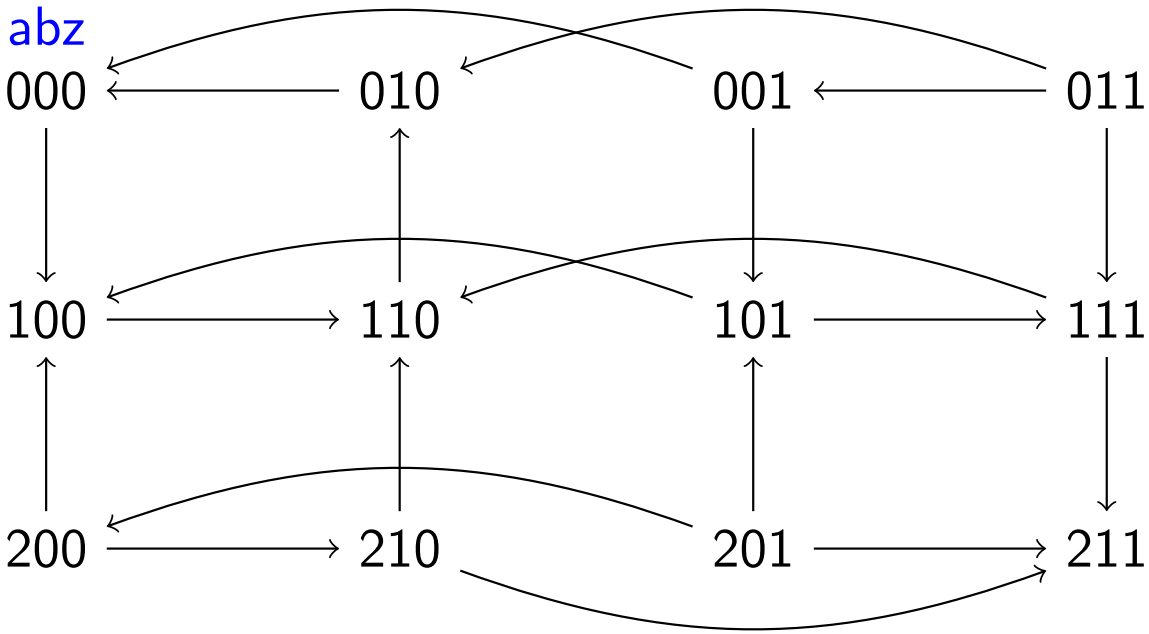
State-graph

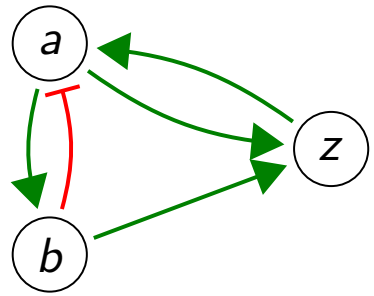
The state-graph depicts explicitly the whole dynamics



State-graph

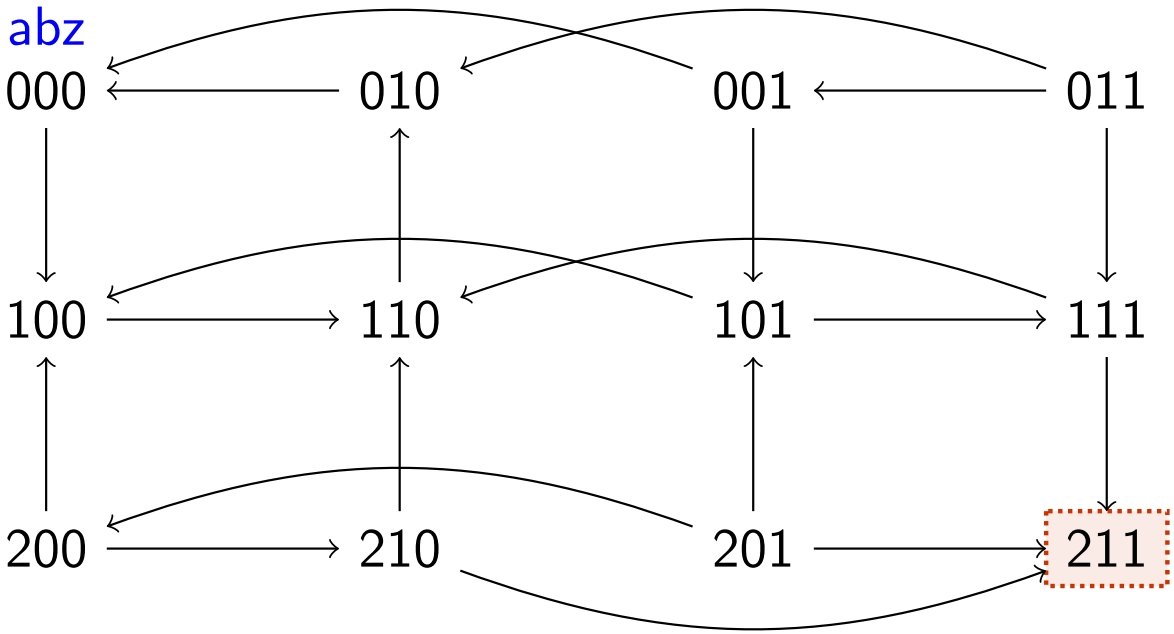
The state-graph depicts explicitly the whole dynamics

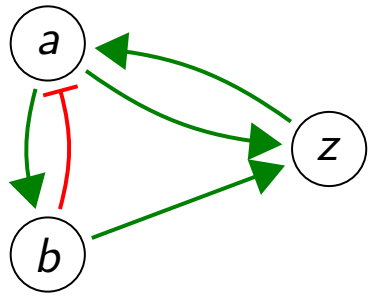




State-graph

The state-graph depicts explicitly the whole dynamics

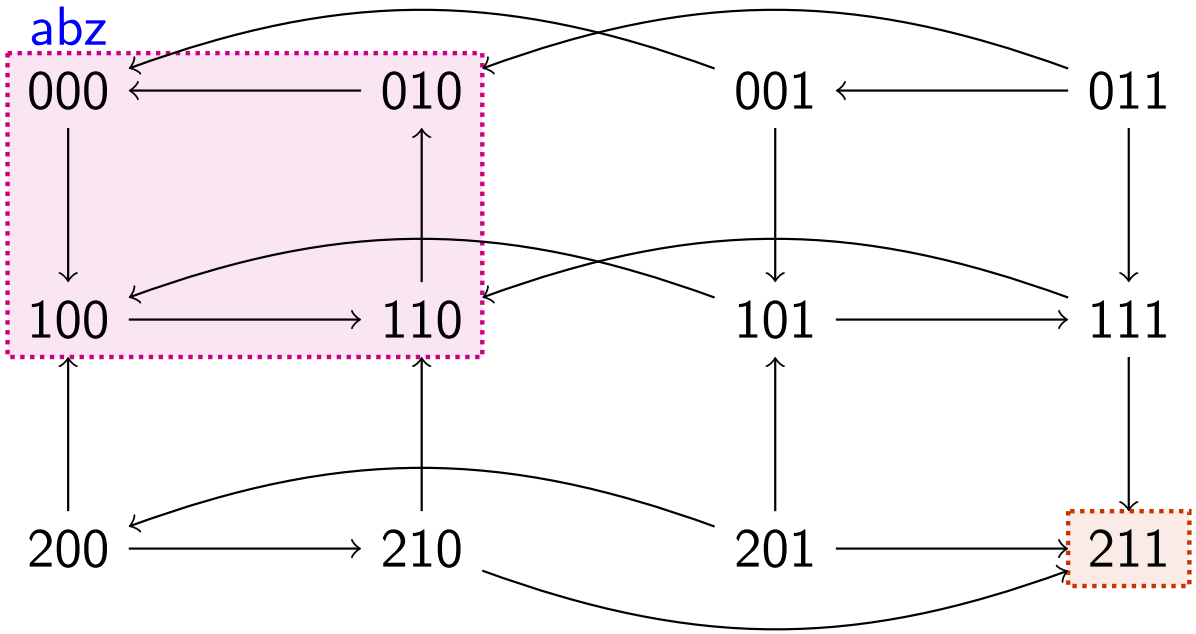


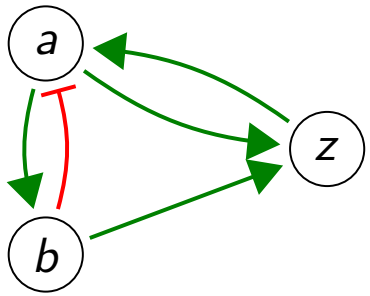


- **Stable state** = state with no successors

State-graph

The state-graph depicts explicitly the whole dynamics

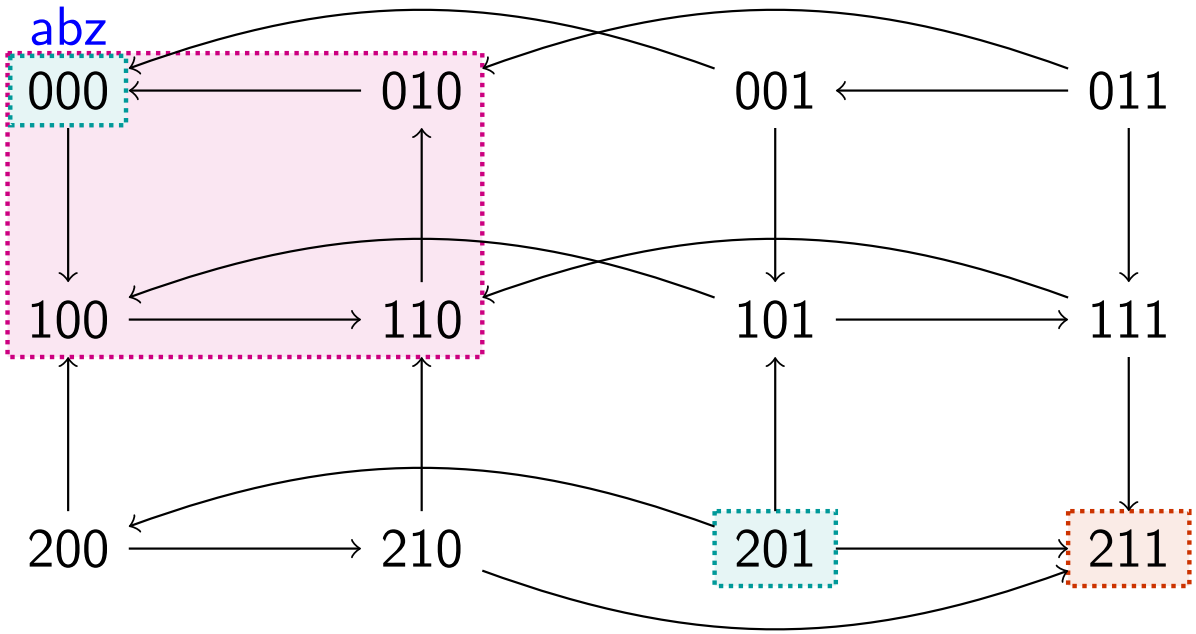


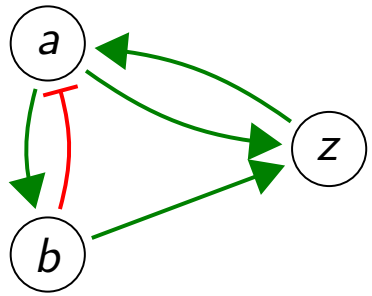


- **Stable state** = state with no successors
- **Complex attractor** = minimal loop or composition of loops from which the dynamics cannot escape

State-graph

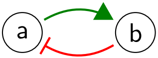
The state-graph depicts explicitly the whole dynamics



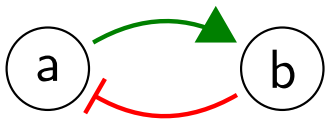
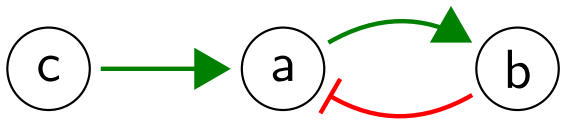


- **Stable state** = state with no successors
- **Complex attractor** = minimal loop or composition of loops from which the dynamics cannot escape
- **Reachability** = from **201**, can I reach **000**?

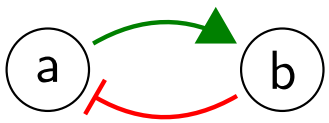
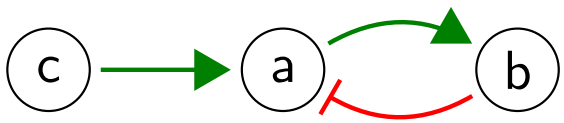
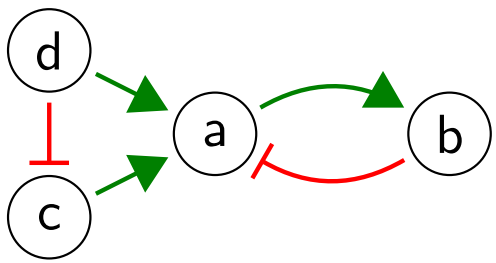
Combinatorial explosion

Model	Possible states
	4

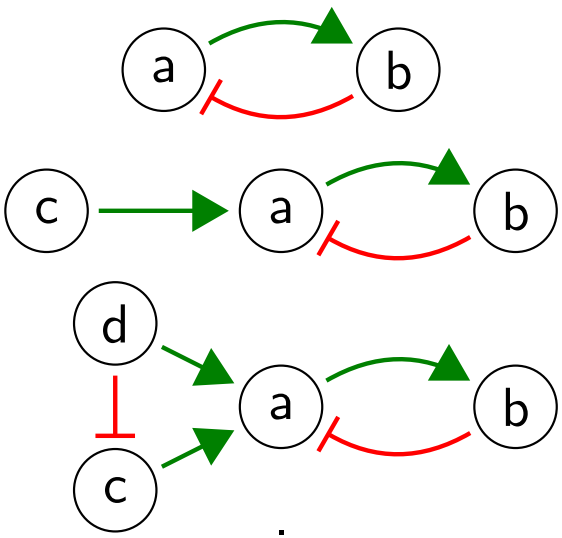
Combinatorial explosion

Model	Possible states
	4
	8

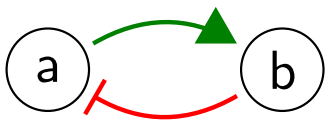
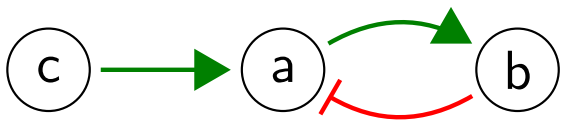
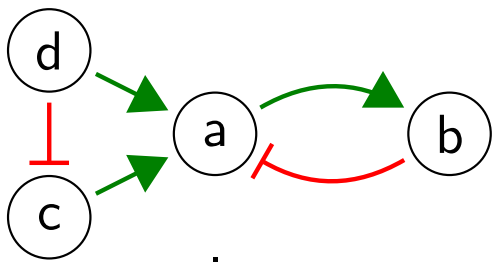
Combinatorial explosion

Model	Possible states
	4
	8
	16

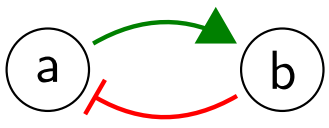
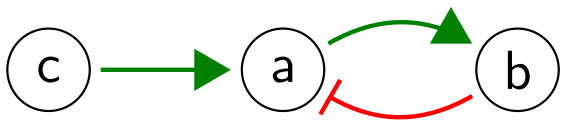
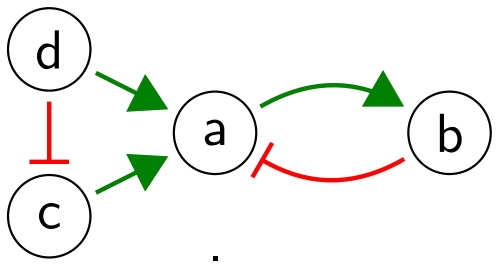
Combinatorial explosion

Model	Possible states
	4
	8
	16
⋮	⋮
(10)	1024

Combinatorial explosion

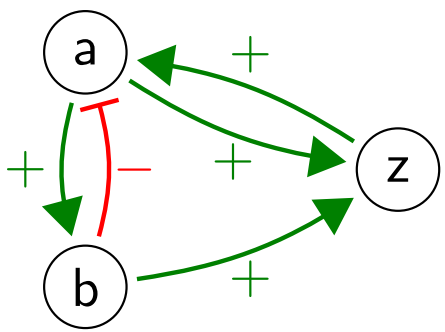
Model	Possible states
	4
	8
	16
⋮	⋮
(10)	1024
(20)	1048576

Combinatorial explosion

Model	Possible states
	4
	8
	16
⋮	⋮
(10)	1024
(20)	1048576
(100)	1267650600000000000000000000

Conjectures of René Thomas

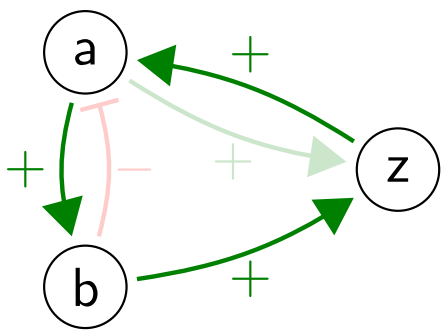
[Thomas, *Num. Methods in the Study of Crit. Phenomena*, 1981]



Conjectures of René Thomas

[Thomas, *Num. Methods in the Study of Crit. Phenomena*, 1981]

- Multiple **stable states** ⇒ positive cycle



110

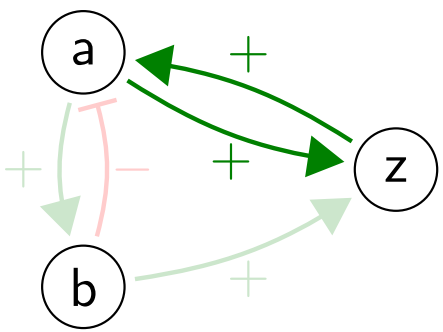
111

000

Conjectures of René Thomas

[Thomas, *Num. Methods in the Study of Crit. Phenomena*, 1981]

- Multiple **stable states** ⇒ positive cycle



110

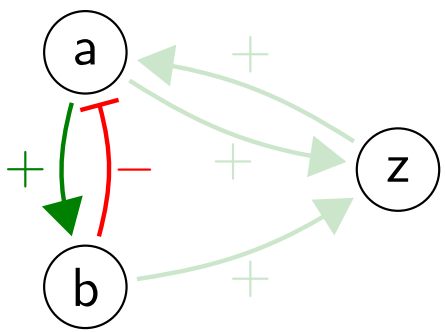
111

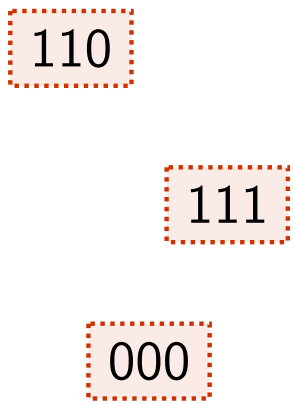
000

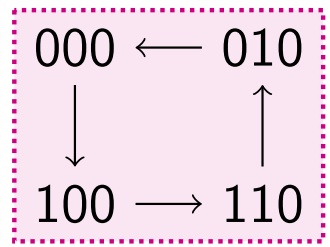
Conjectures of René Thomas

[Thomas, *Num. Methods in the Study of Crit. Phenomena*, 1981]

- Multiple **stable states** ⇒ positive cycle
- Sustained oscillations (**complex attractor**) ⇒ negative cycle



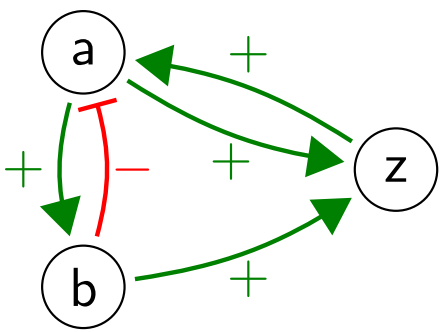


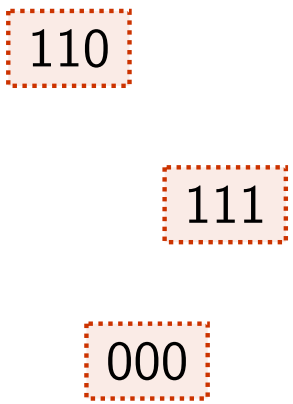


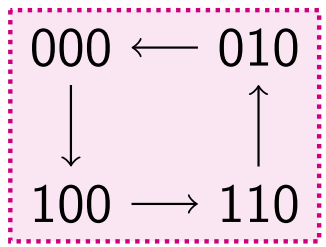
Conjectures of René Thomas

[Thomas, *Num. Methods in the Study of Crit. Phenomena*, 1981]

- Multiple **stable states** ⇒ positive cycle
- Sustained oscillations (**complex attractor**) ⇒ negative cycle





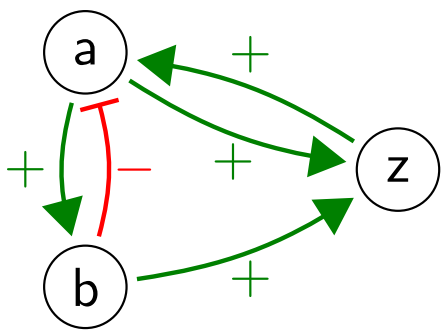


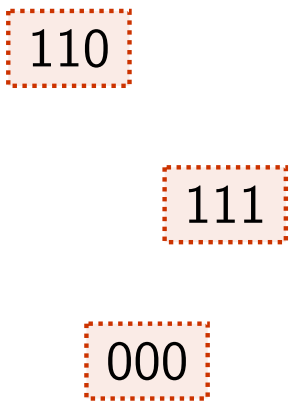
Proofs: [Remy *et al.*, *Advances in Applied Mathematics*, 2008]
[Richard, *Advances in Applied Mathematics*, 2010]
[Richard & Comet, *Discrete Applied Mathematics*, 2007]

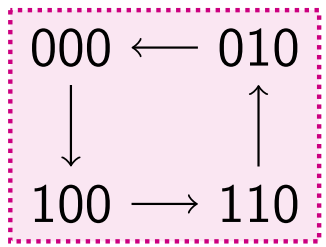
Conjectures of René Thomas

[Thomas, *Num. Methods in the Study of Crit. Phenomena*, 1981]

- Multiple **stable states** ⇒ positive cycle
  No positive cycle ⇒ The stable state (if any) is unique
- Sustained oscillations (**complex attractor**) ⇒ negative cycle





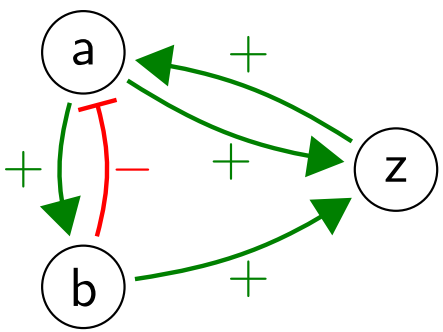


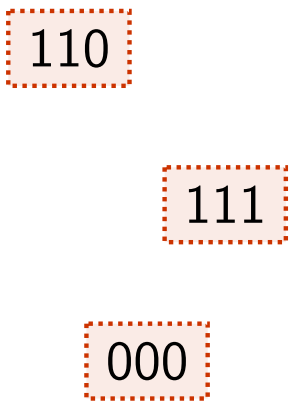
Proofs: [Remy *et al.*, *Advances in Applied Mathematics*, 2008]
[Richard, *Advances in Applied Mathematics*, 2010]
[Richard & Comet, *Discrete Applied Mathematics*, 2007]

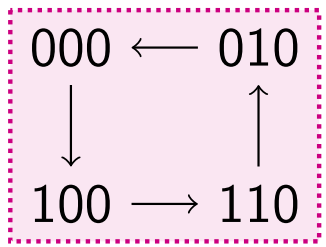
Conjectures of René Thomas

[Thomas, *Num. Methods in the Study of Crit. Phenomena*, 1981]

- Multiple **stable states** ⇒ positive cycle
  No positive cycle ⇒ The stable state (if any) is unique
- Sustained oscillations (**complex attractor**) ⇒ negative cycle
  No negative cycle ⇒ No complex attractor (only stable states)



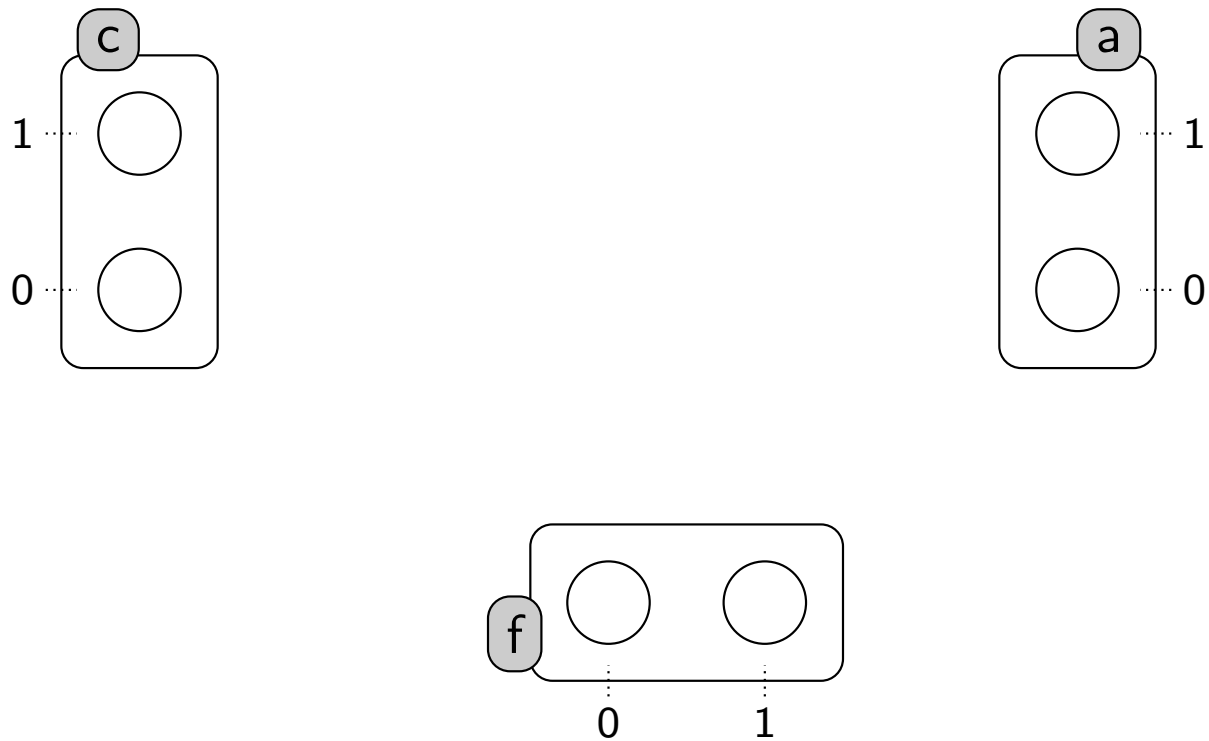




Proofs: [Remy *et al.*, *Advances in Applied Mathematics*, 2008]
[Richard, *Advances in Applied Mathematics*, 2010]
[Richard & Comet, *Discrete Applied Mathematics*, 2007]

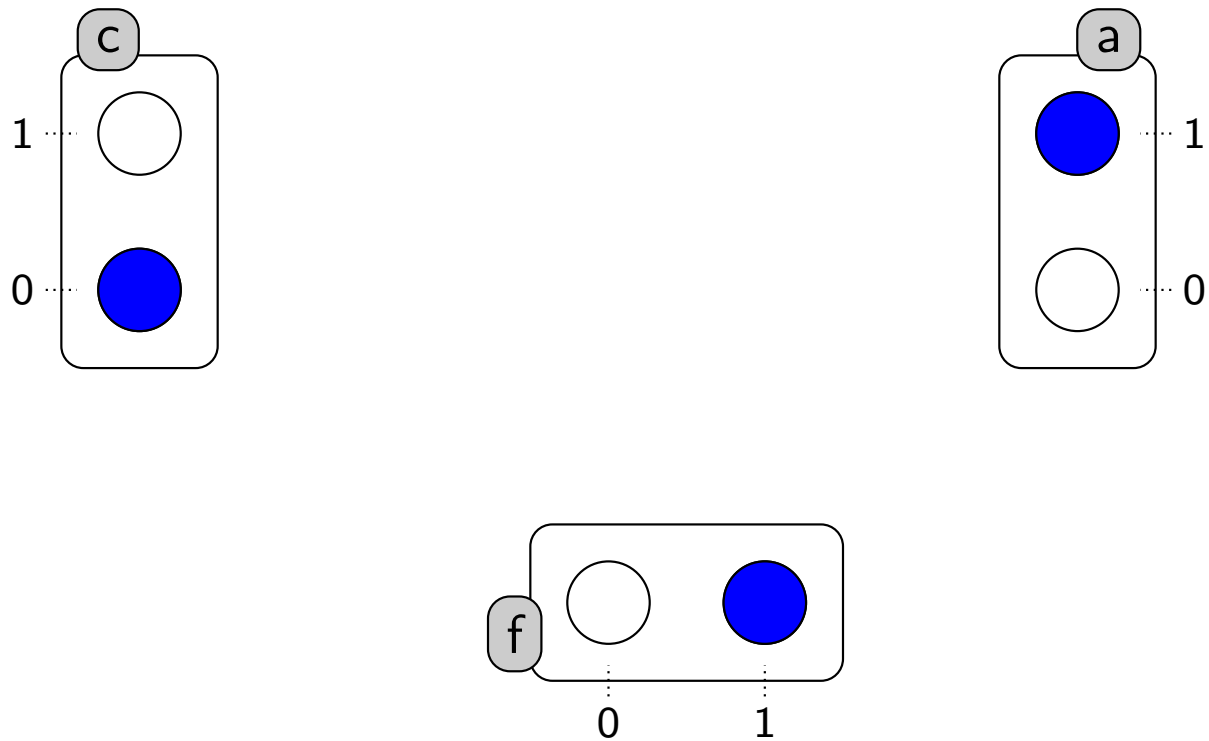
Asynchronous Automata Networks (AAN)

[Paulevé *et al.*, *Transactions on Computational Systems Biology*, 2011]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



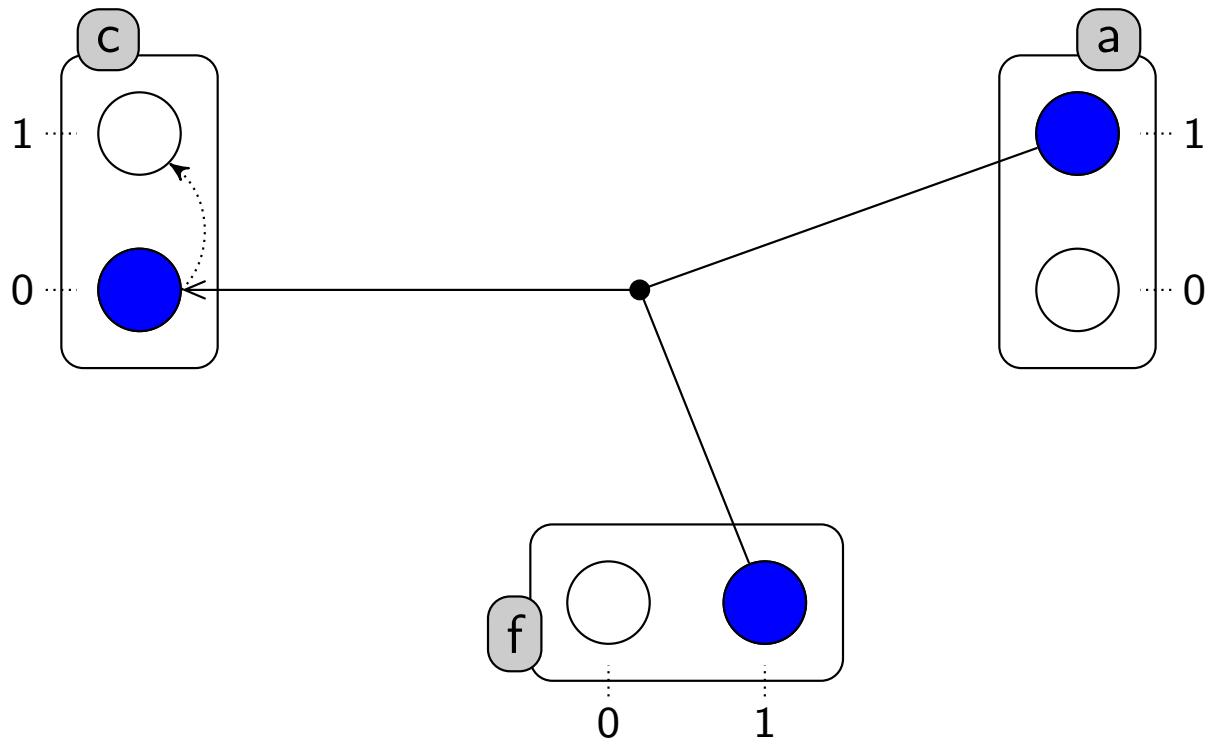
Asynchronous Automata Networks (AAN)

[Paulevé *et al.*, *Transactions on Computational Systems Biology*, 2011]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



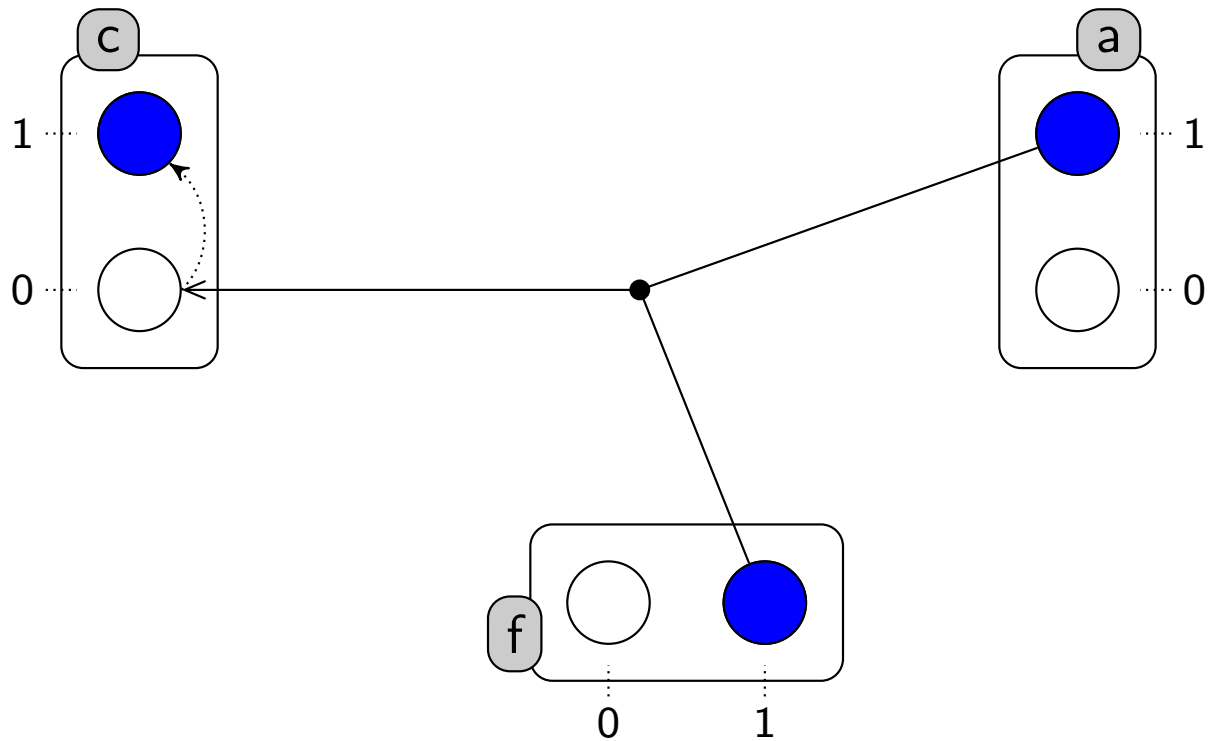
Asynchronous Automata Networks (AAN)

[Paulevé *et al.*, *Transactions on Computational Systems Biology*, 2011]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



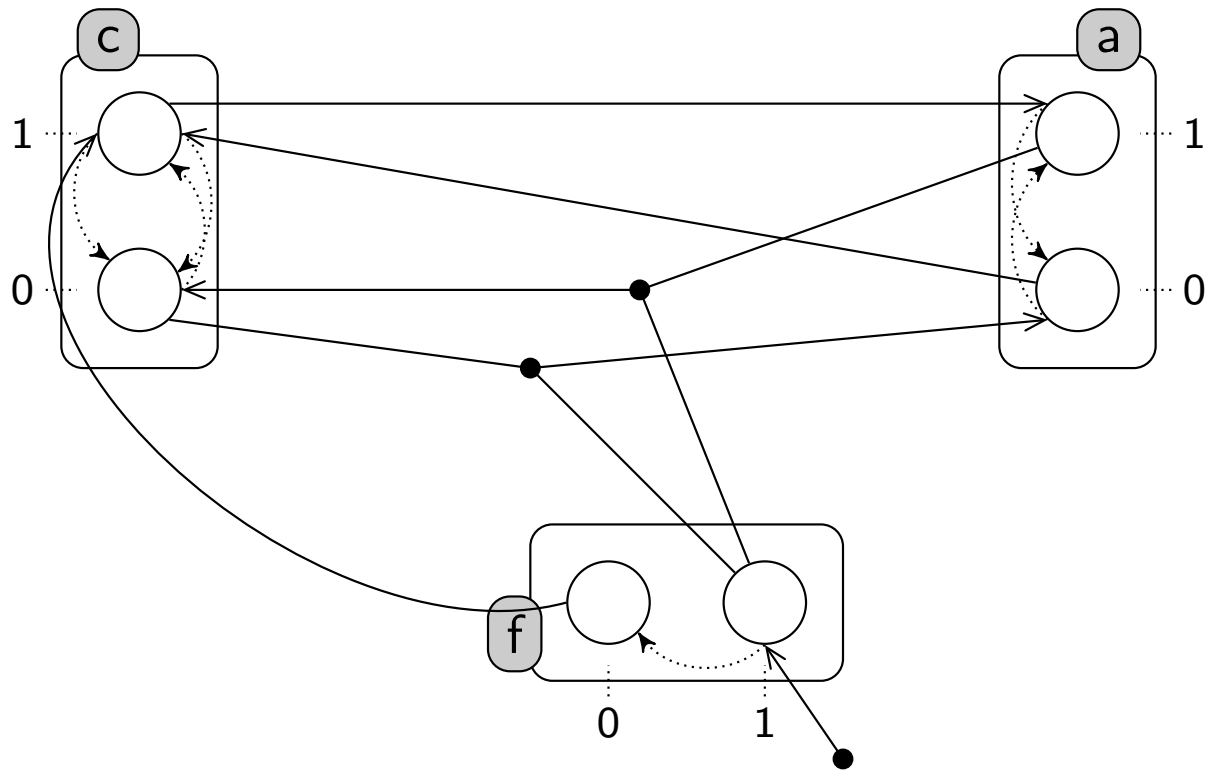
Asynchronous Automata Networks (AAN)

[Paulevé *et al.*, *Transactions on Computational Systems Biology*, 2011]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



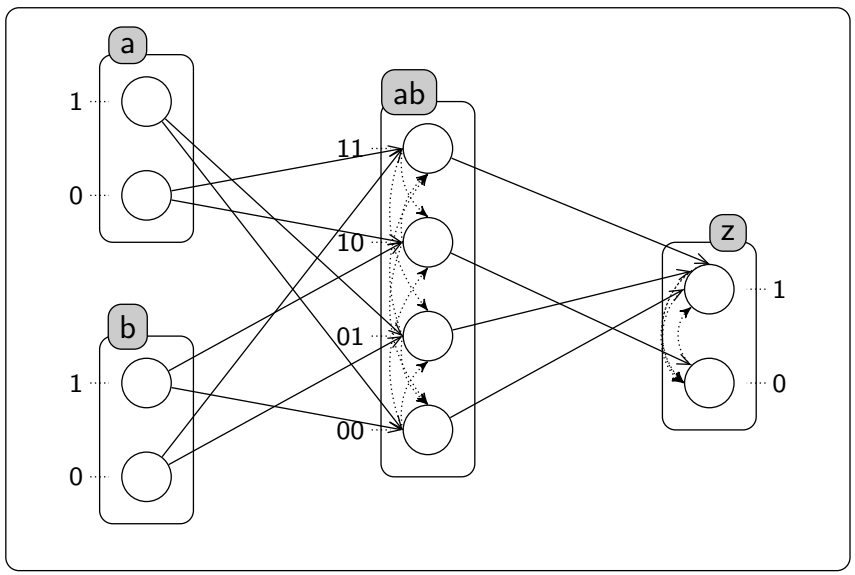
Asynchronous Automata Networks (AAN)

[Paulevé *et al.*, *Transactions on Computational Systems Biology*, 2011]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]
Model from [François *et al.*, *Molecular Systems Biology*, 2007]



Translation of AAN models

[Folschette *et al.*, *Computational Methods in Systems Biology*, 2012]

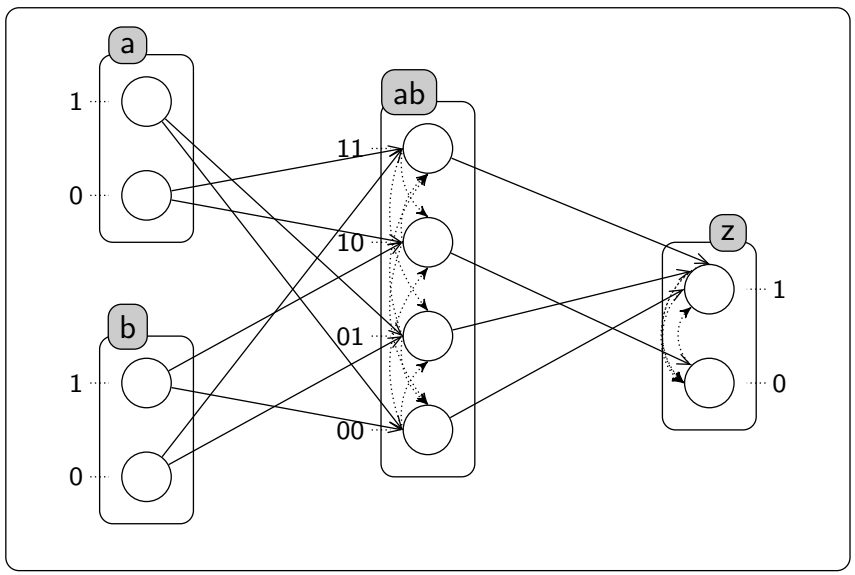


Process Hitting

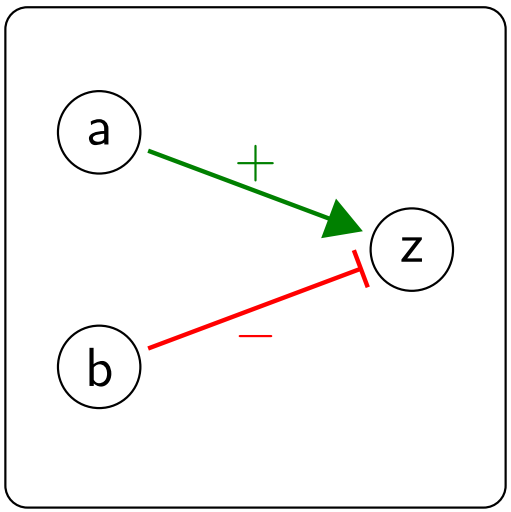
Efficient but recent

Translation of AAN models

[Folschette *et al.*, *Computational Methods in Systems Biology*, 2012]



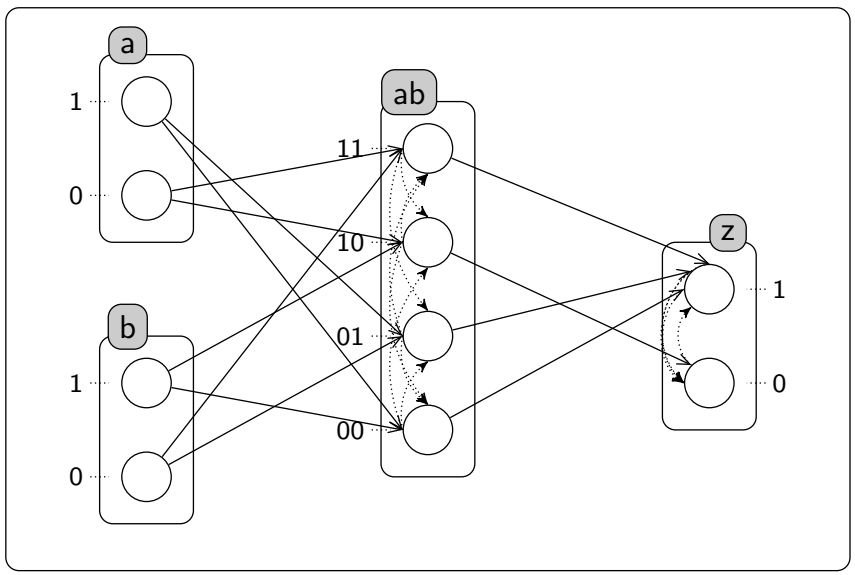
Process Hitting
Efficient but recent



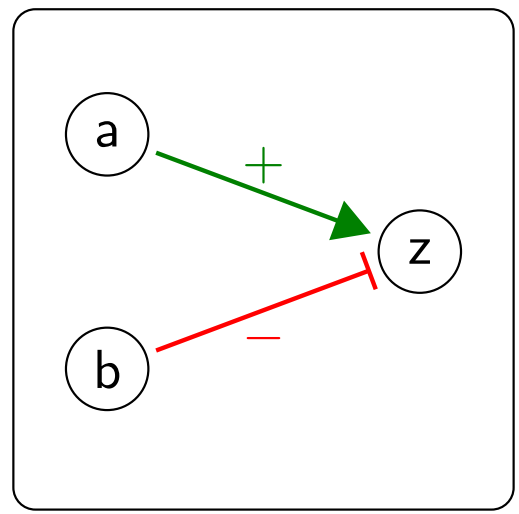
Thomas modeling
Widespread & readable

Translation of AAN models

[Folschette *et al.*, *Computational Methods in Systems Biology*, 2012]



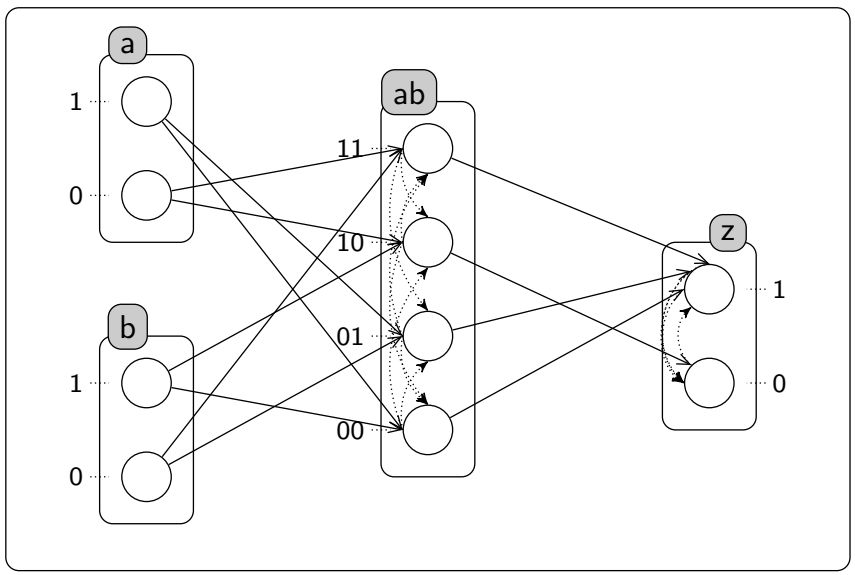
Process Hitting
Efficient but recent

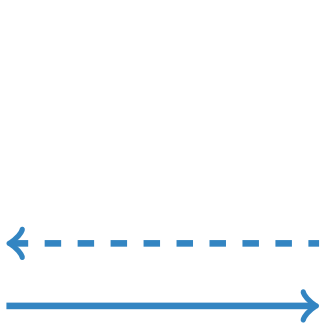


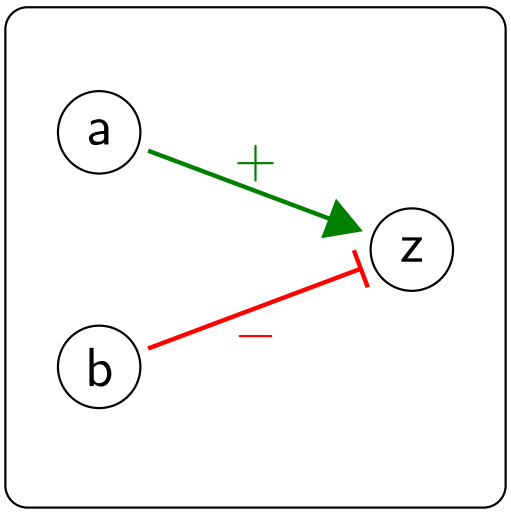
Thomas modeling
Widespread & readable

Translation of AAN models

[Folschette *et al.*, *Computational Methods in Systems Biology*, 2012]





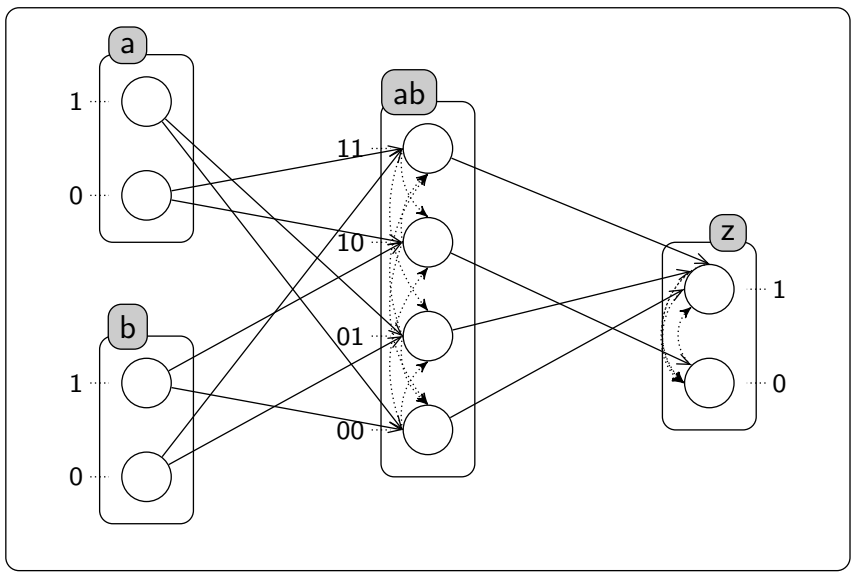


Process Hitting
Efficient but recent

Thomas modeling
Widespread & readable

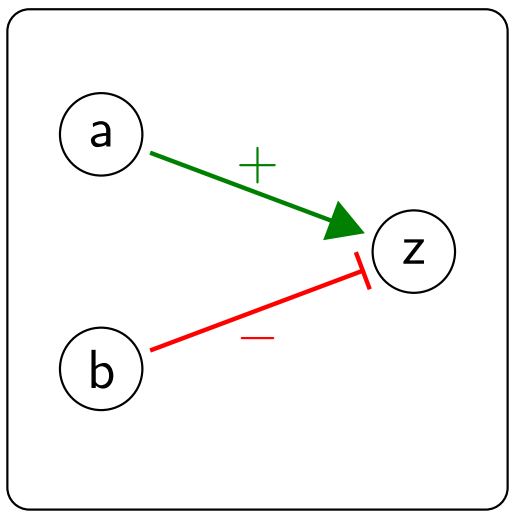
Towards AANs

[Folschette *et al.*, *CS2Bio*, 2013]



Process Hitting
Loose behavior

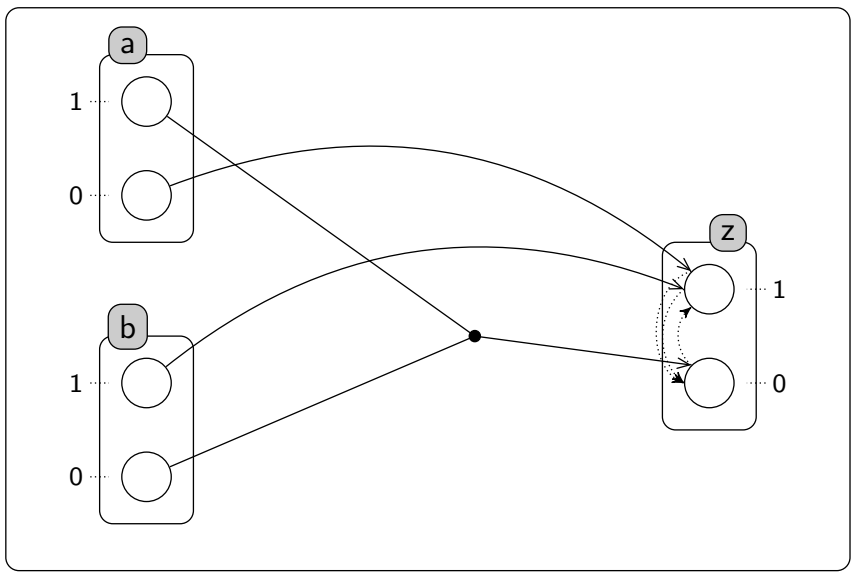
≠



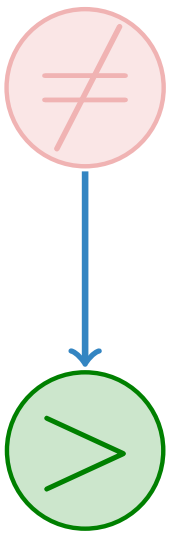
Thomas modeling
Expected behavior

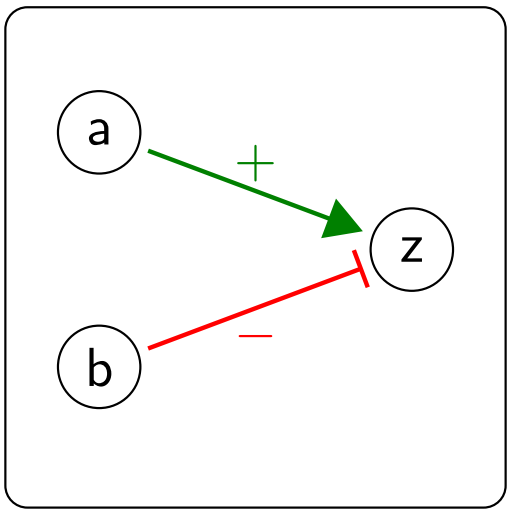
Towards AANs

[Folschette *et al.*, *CS2Bio*, 2013]



AANs
Accurate behavior

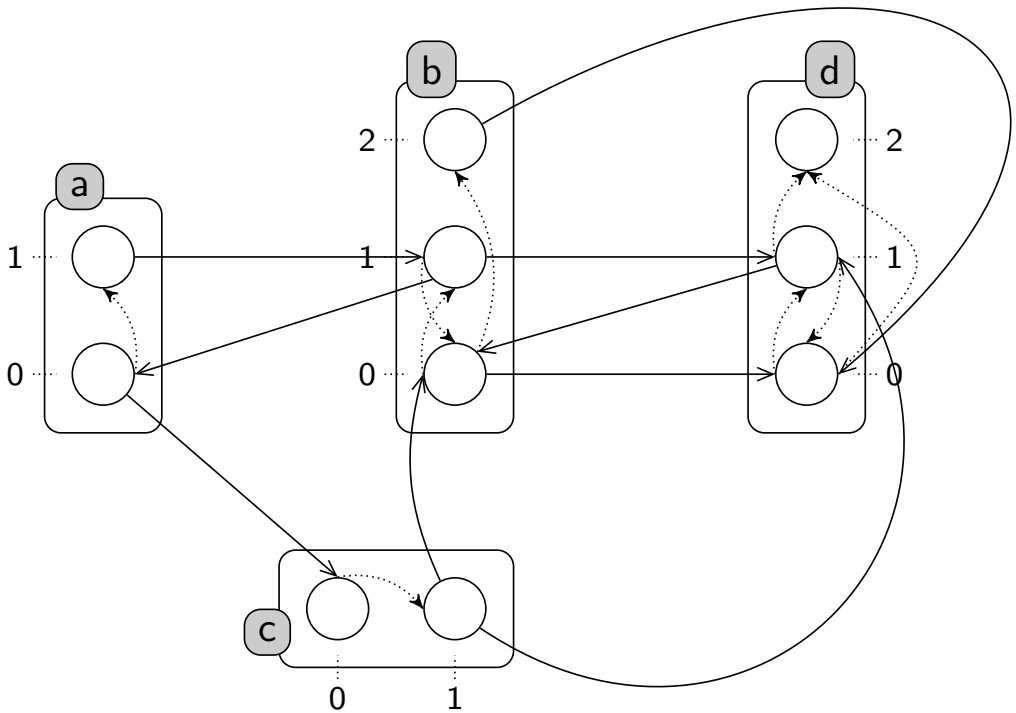




Thomas modeling
Expected behavior

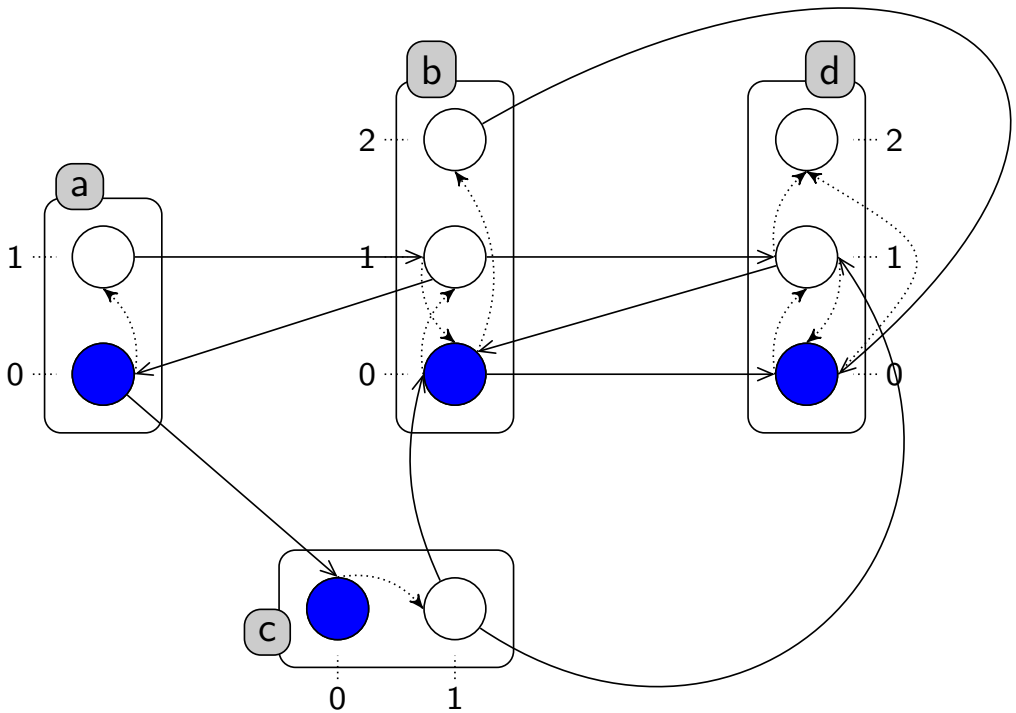
Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

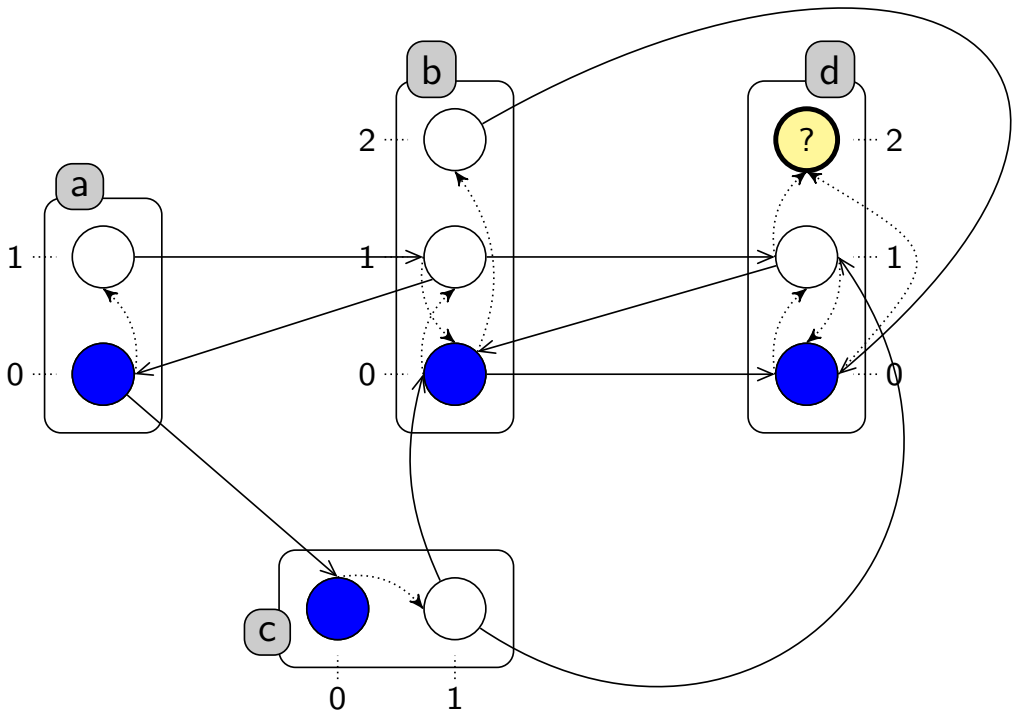


- Initial state

$\langle a_1, b_0, c_0, d_0 \rangle$

Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



- Initial state

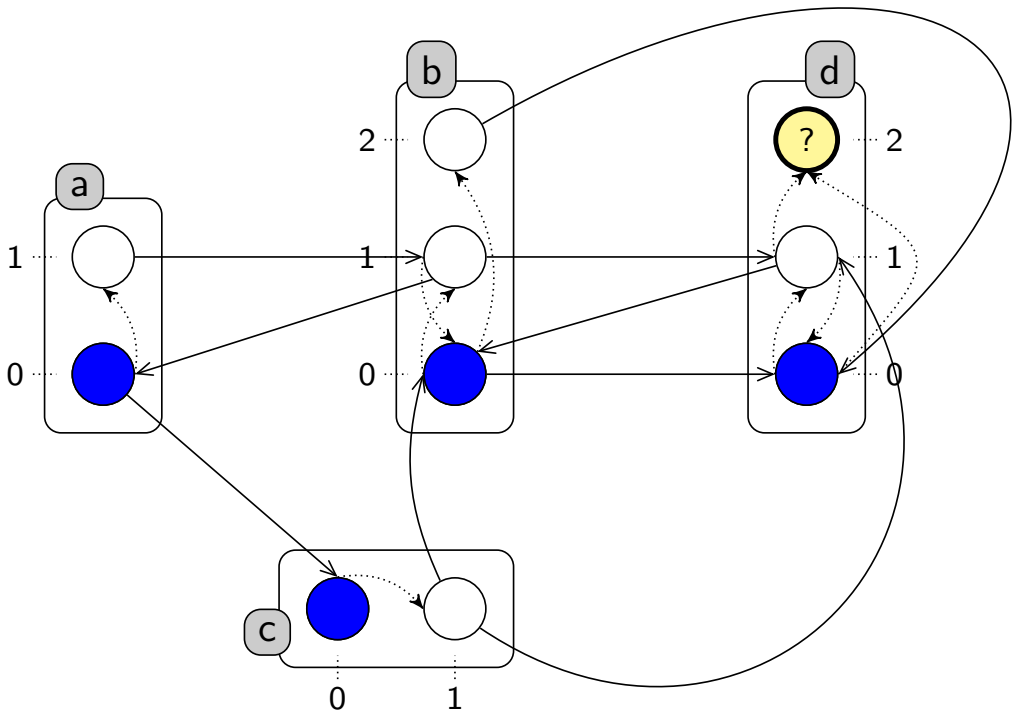
$$\langle a_1, b_0, c_0, d_0 \rangle$$

- Objective(s)

$$[\ \Rsh\ d_2\ ]$$

Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



- Initial state

$\langle a_1, b_0, c_0, d_0 \rangle$

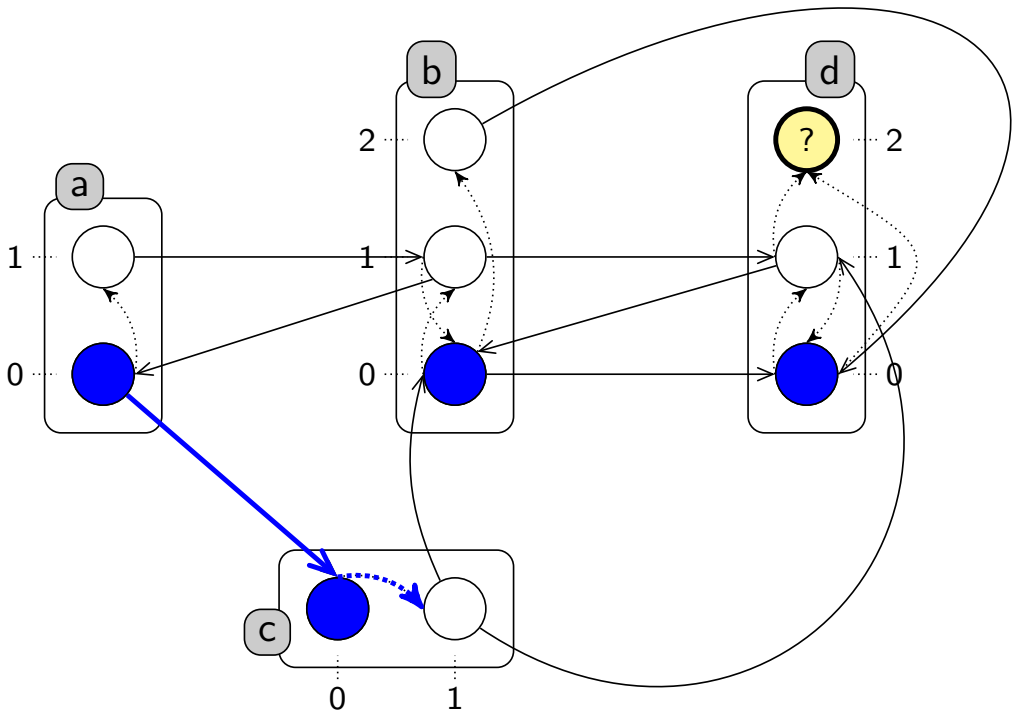
- Objective(s)

$[\ \Rsh\ d_2\ ]$

$\rightarrow$ Sequence of actions?

Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



- Initial state

$\langle a_1, b_0, c_0, d_0 \rangle$

- Objective(s)

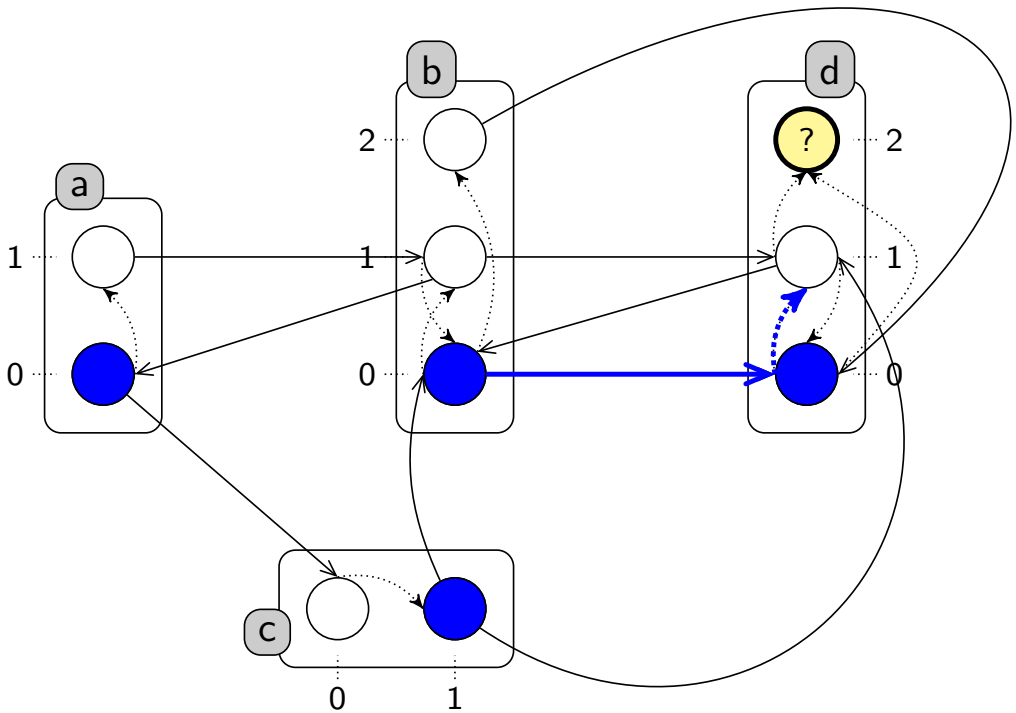
$[\ \Rsh\ d_2\ ]$

$\rightarrow$ Sequence of actions?

$a_0 \rightarrow c_0 \Rsh c_1$

Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



- Initial state

$\langle a_1, b_0, c_0, d_0 \rangle$

- Objective(s)

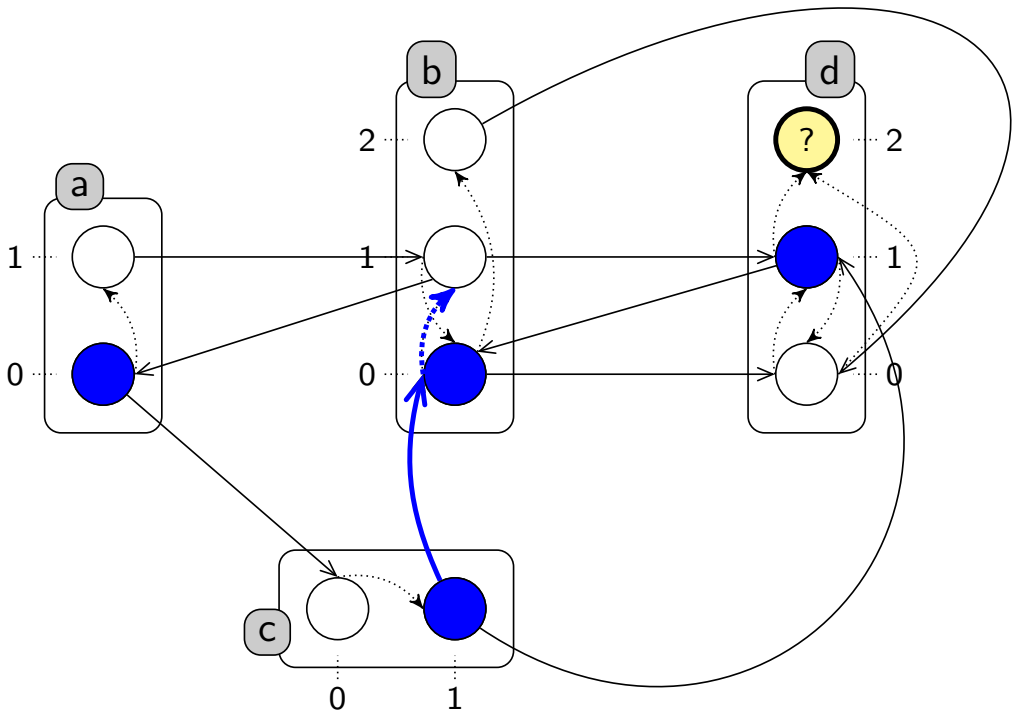
$[\ \Rsh\ d_2\ ]$

$\rightarrow$ Sequence of actions?

$a_0 \rightarrow c_0 \Rsh c_1 :: b_0 \rightarrow d_0 \Rsh d_1$

Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



- Initial state

$\langle a_1, b_0, c_0, d_0 \rangle$

- Objective(s)

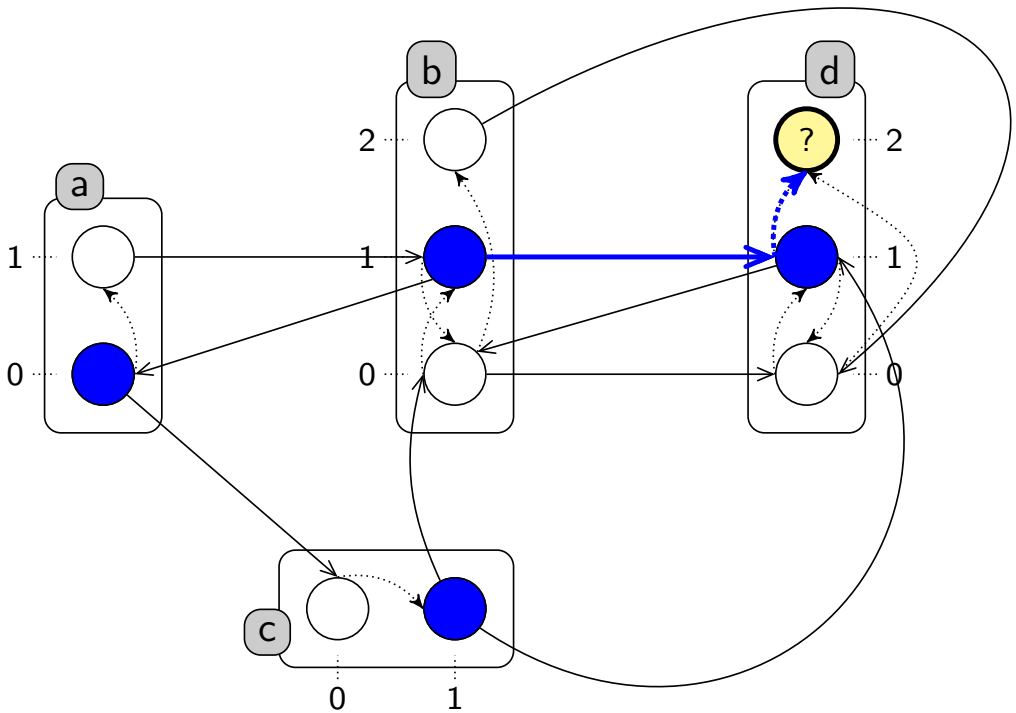
$[\ \Rsh\ d_2\ ]$

$\rightarrow$ Sequence of actions?

$a_0 \rightarrow c_0 \Rsh c_1 :: b_0 \rightarrow d_0 \Rsh d_1 :: c_1 \rightarrow b_0 \Rsh b_1$

Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



- Initial state

$$\langle a_1, b_0, c_0, d_0 \rangle$$

- Objective(s)

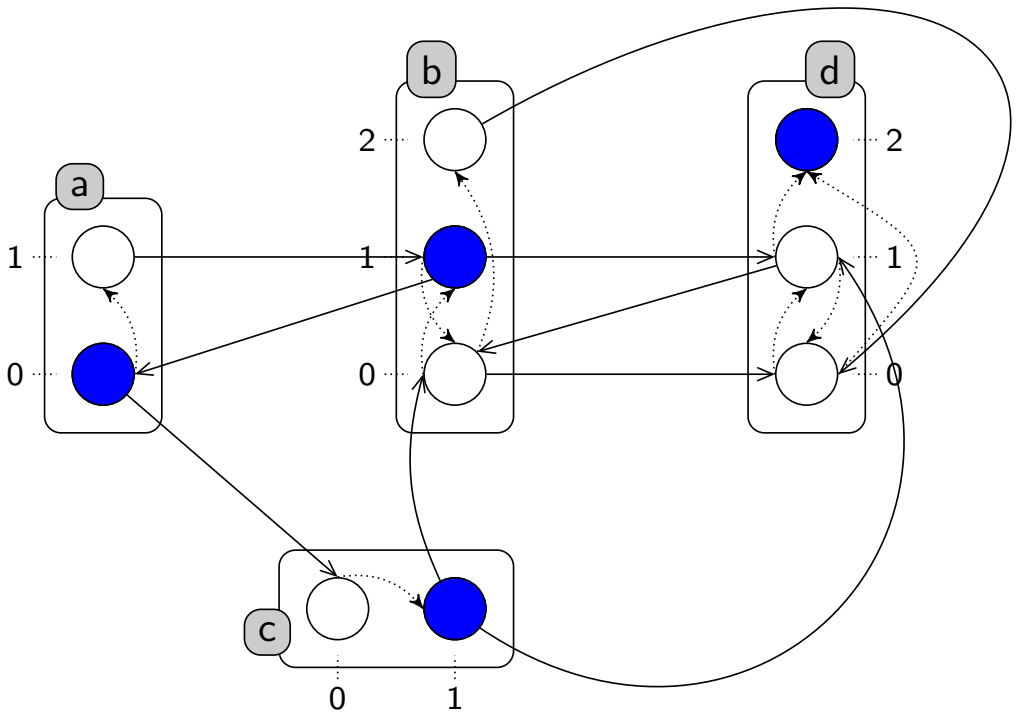
$$[\ \Rsh\ d_2\ ]$$

$\rightarrow$ Sequence of actions?

$$a_0 \rightarrow c_0 \Rsh c_1 :: b_0 \rightarrow d_0 \Rsh d_1 :: c_1 \rightarrow b_0 \Rsh b_1 :: b_1 \rightarrow d_1 \Rsh d_2$$

Static analysis: successive reachability

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]



- Initial state

$$\langle a_1, b_0, c_0, d_0 \rangle$$

- Objective(s)

$$[\ \Rsh\ d_2\ ]$$

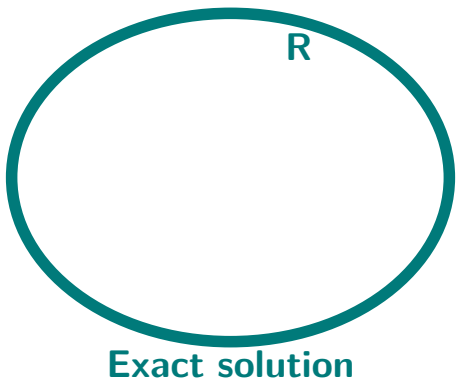
$\rightarrow$ Sequence of actions?

$$a_0 \rightarrow c_0 \Rsh c_1 :: b_0 \rightarrow d_0 \Rsh d_1 :: c_1 \rightarrow b_0 \Rsh b_1 :: b_1 \rightarrow d_1 \Rsh d_2$$

Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

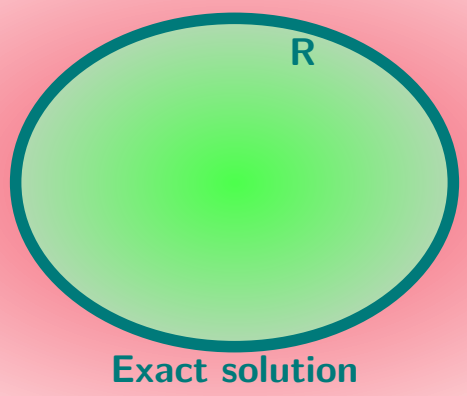
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: $\mathbf{P} \Rightarrow \mathbf{R} \Rightarrow \mathbf{Q}$
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

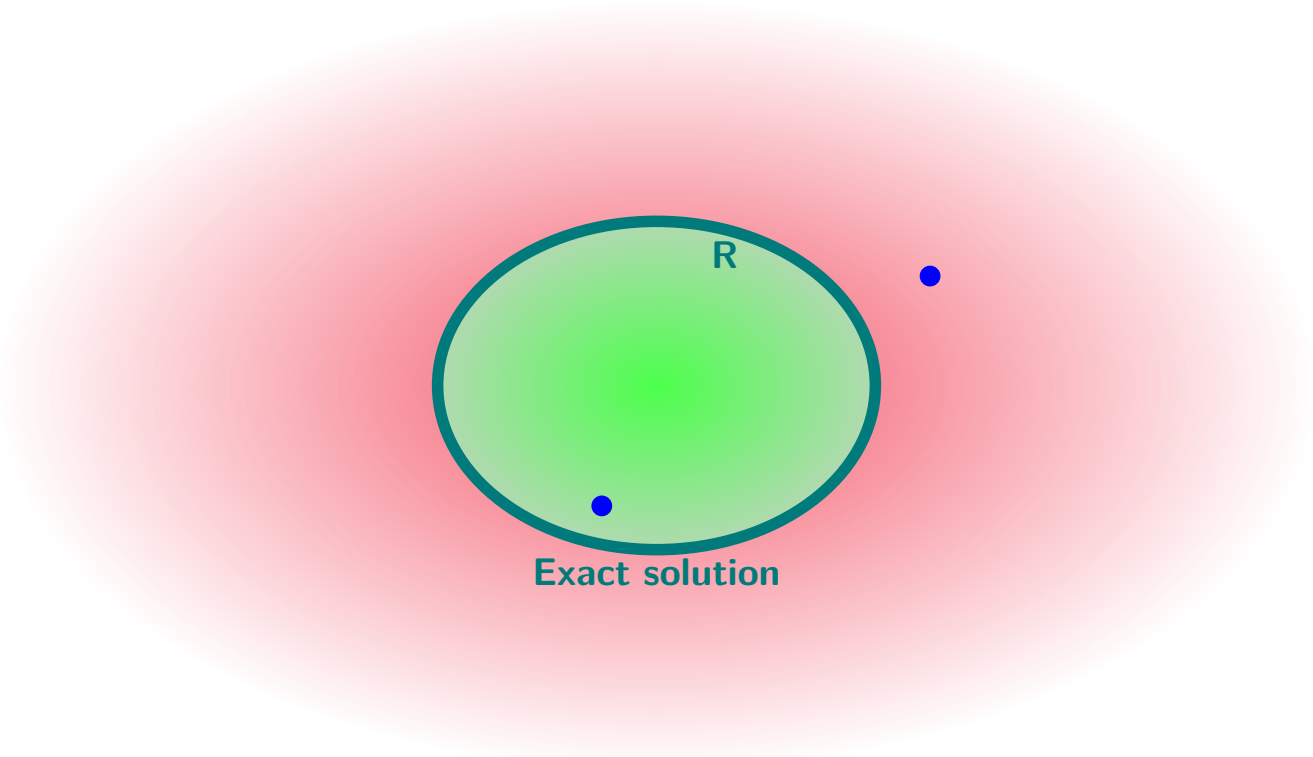
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: $\mathbf{P} \Rightarrow \mathbf{R} \Rightarrow \mathbf{Q}$
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

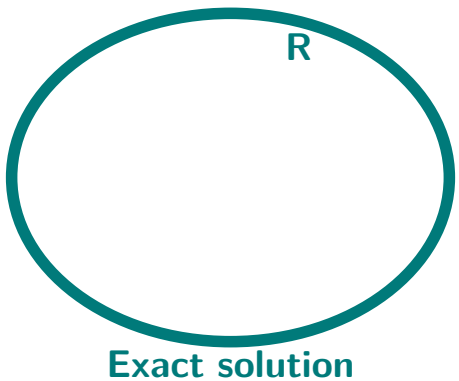
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: **P** $\Rightarrow$ **R** $\Rightarrow$ **Q**
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

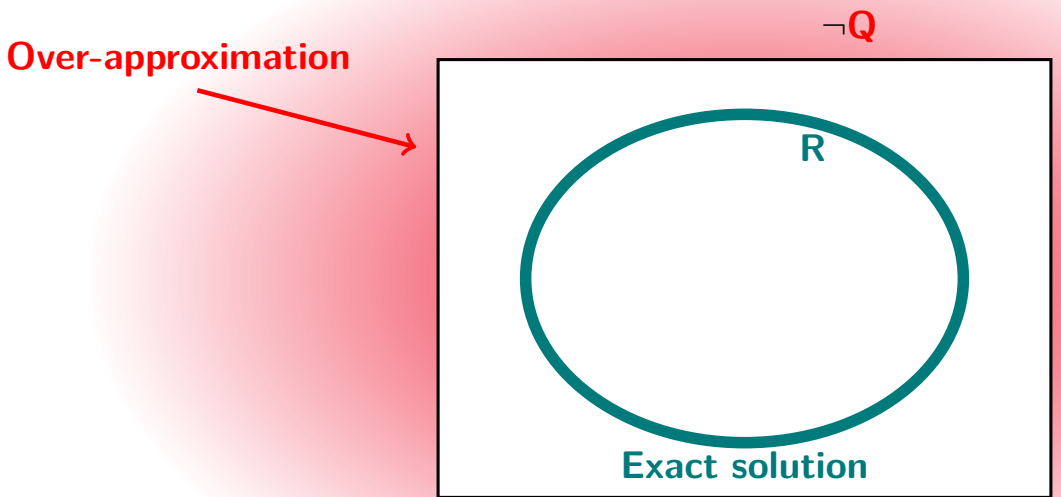
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: $\mathbf{P} \Rightarrow \mathbf{R} \Rightarrow \mathbf{Q}$
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

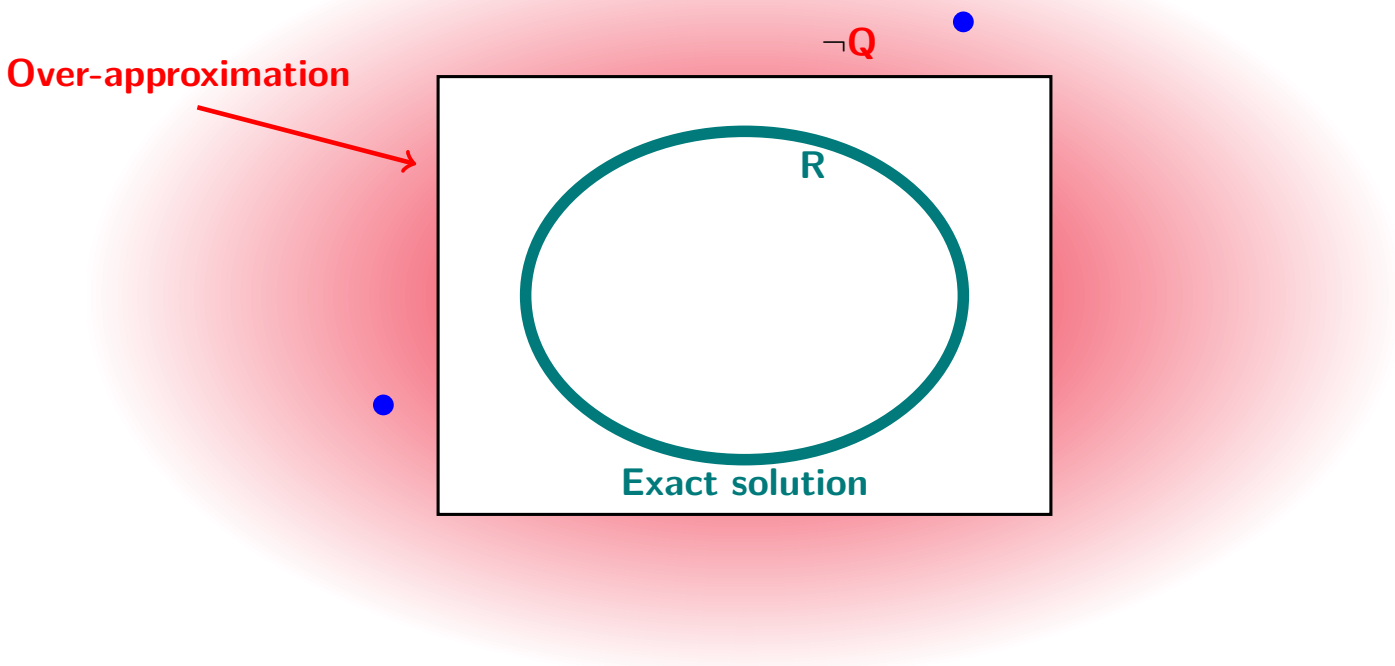
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: **P** $\Rightarrow$ **R** $\Rightarrow$ **Q**
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

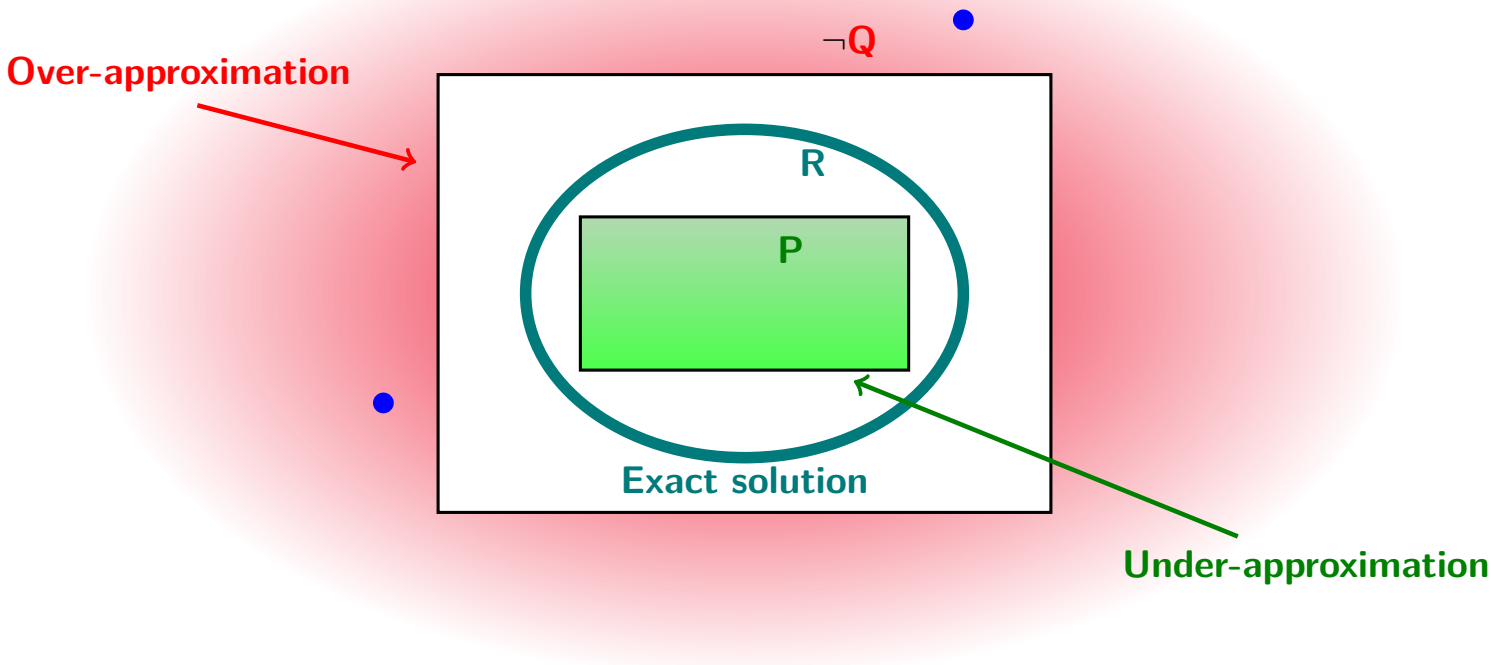
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: **P** $\Rightarrow$ **R** $\Rightarrow$ **Q**
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

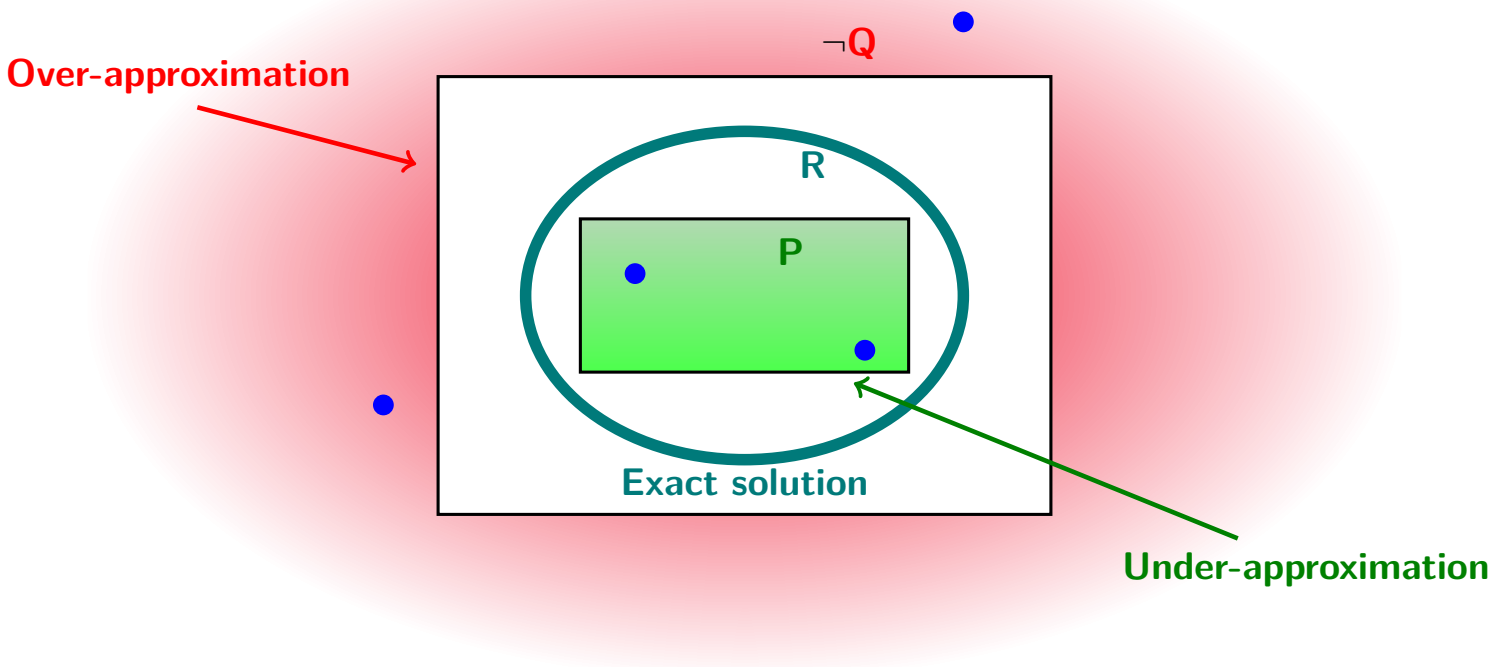
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: **P** ⇒ **R** ⇒ **Q**
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

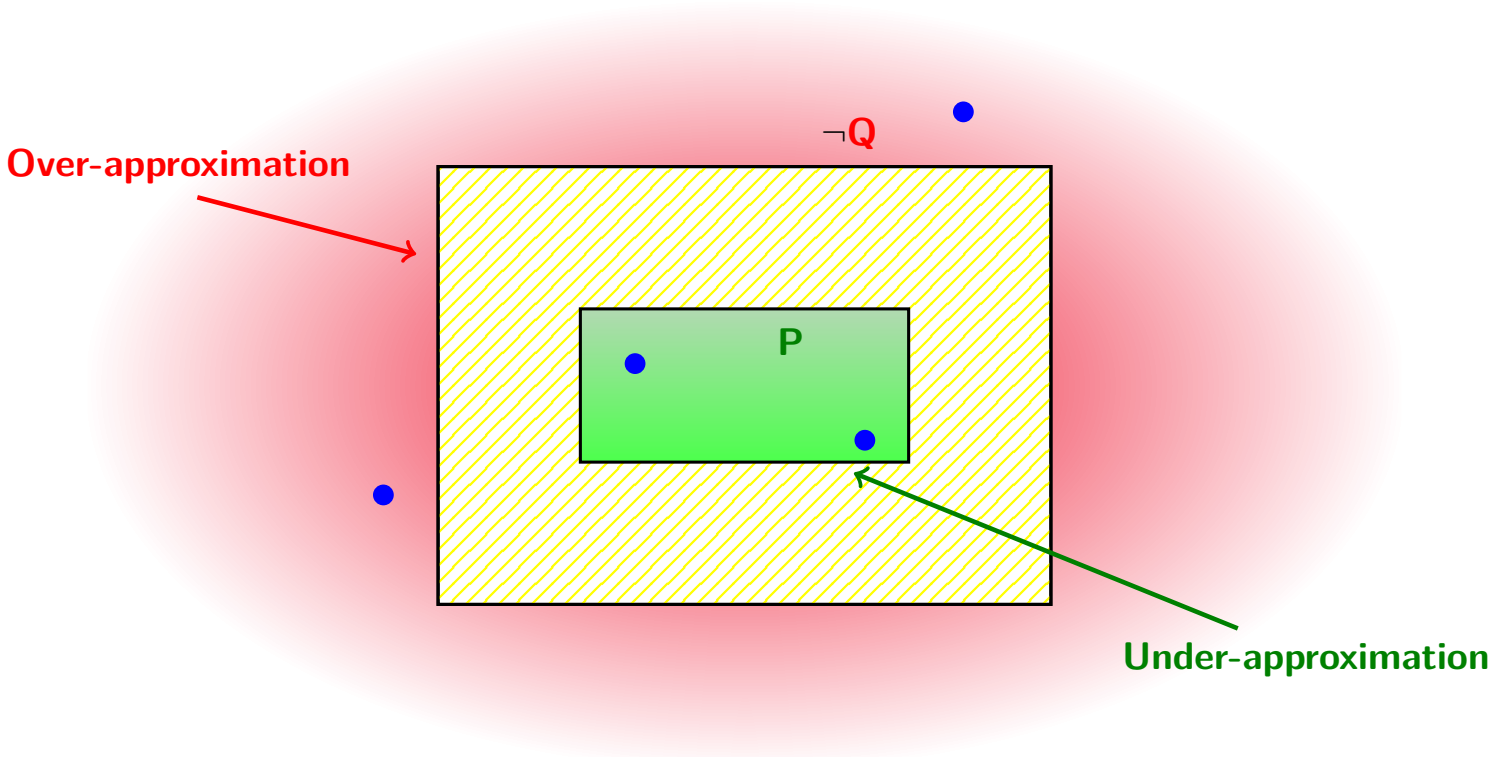
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: **P** ⇒ **R** ⇒ **Q**
  so that computing **P** and **Q** is faster (roughly **polynomial**)



Approximations of the Dynamics

[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

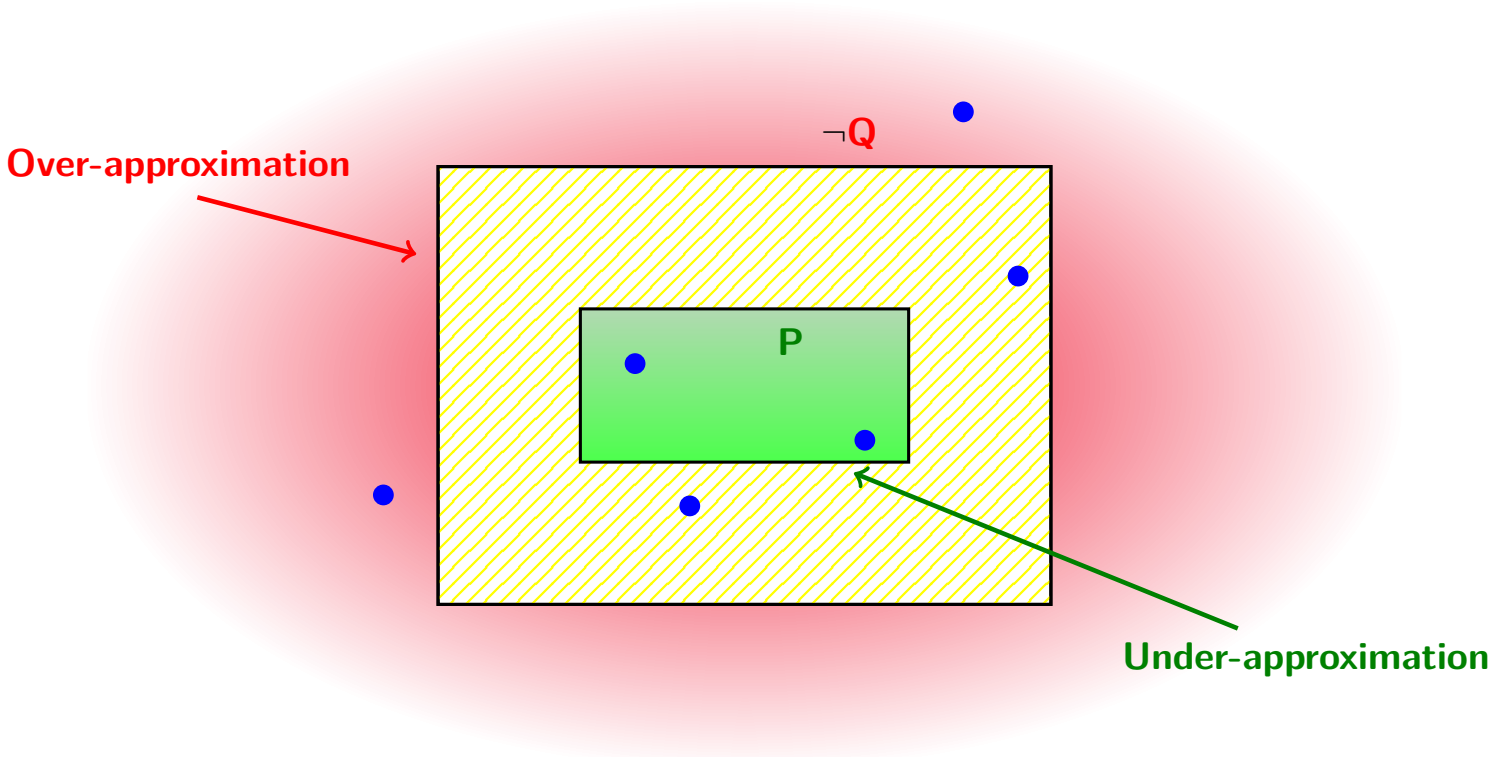
- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: **P** $\Rightarrow$ **R** $\Rightarrow$ **Q**
  so that computing **P** and **Q** is faster (roughly **polynomial**)



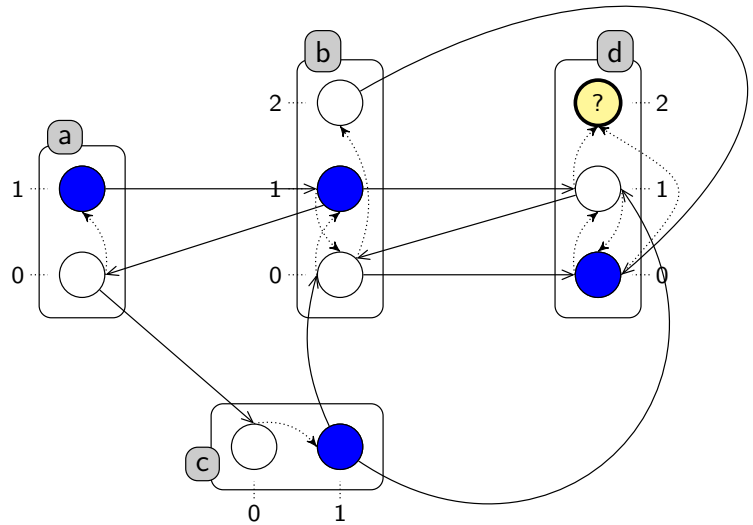
Approximations of the Dynamics

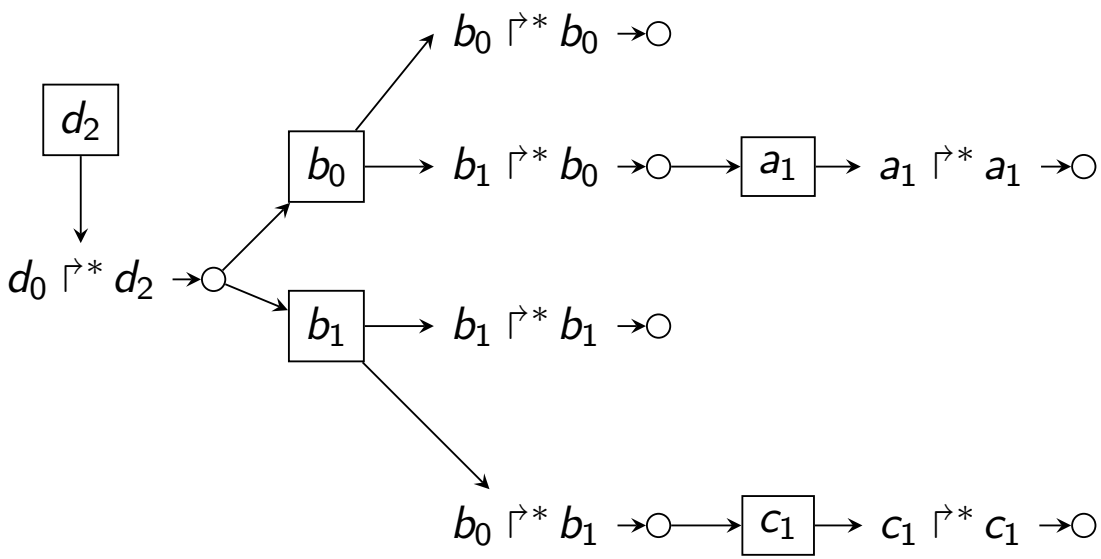
[Paulevé *et al.*, *Mathematical Structures in Computer Science*, 2012]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

- Directly checking **R** is hard (**exponential**)
- Rather check **approximations** **P** and **Q** so that: **P** ⇒ **R** ⇒ **Q**
  so that computing **P** and **Q** is faster (roughly **polynomial**)



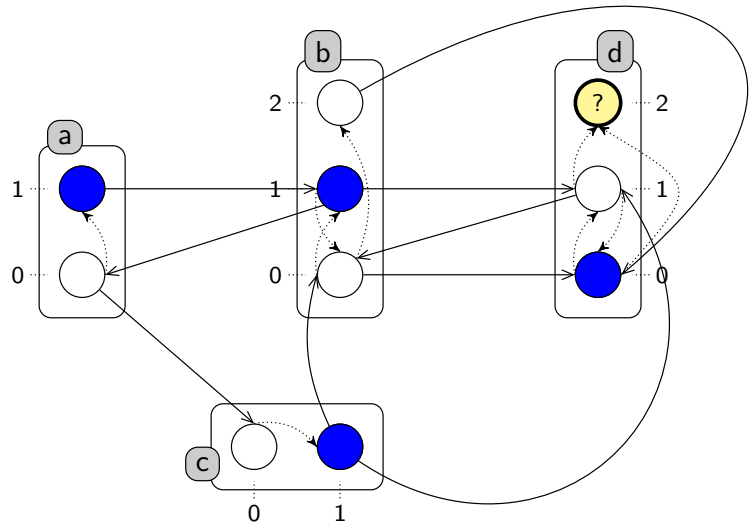
Under-approximation





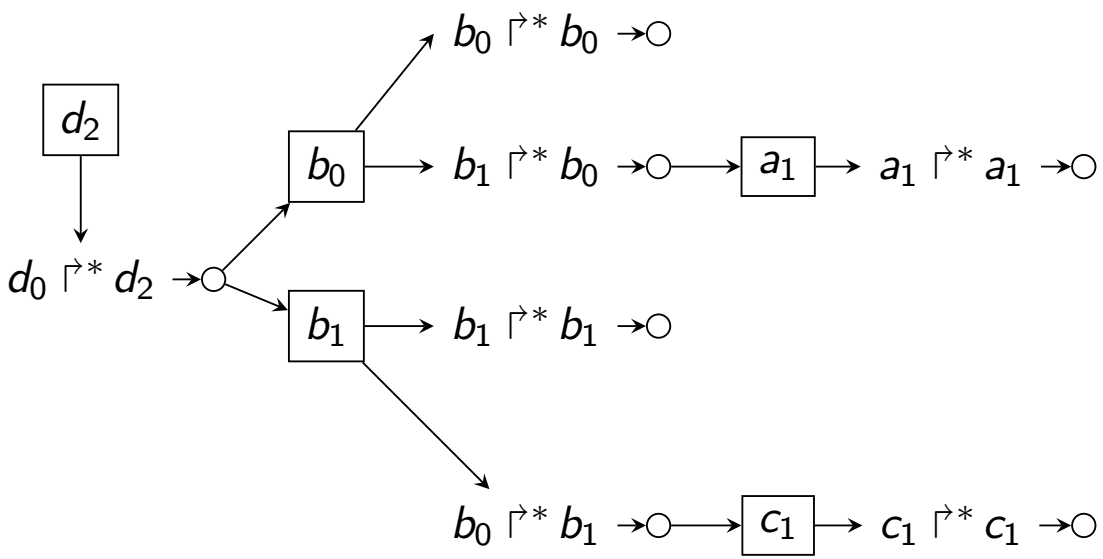
d_2	Required local state
$d_0 \lceil^* d_2$	Objective
○	Solution to an objective

Under-approximation

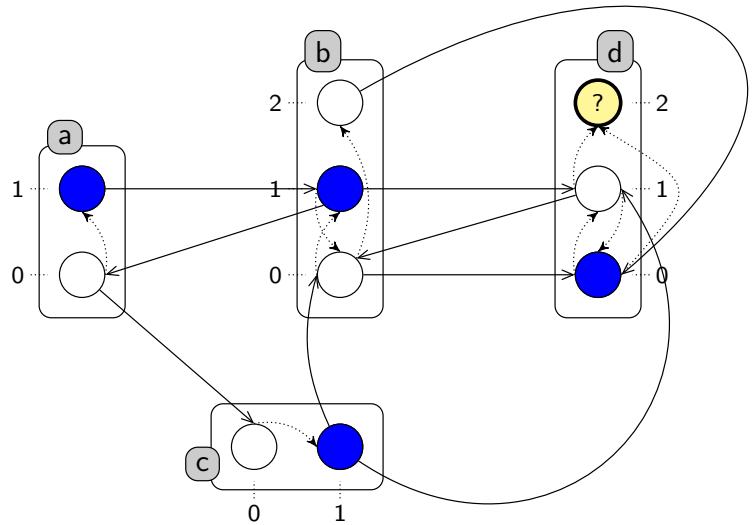


Sufficient condition P:

- no cycle
- each objective has a solution



d_2	Required local state
$d_0 \Rightarrow^* d_2$	Objective
○	Solution to an objective

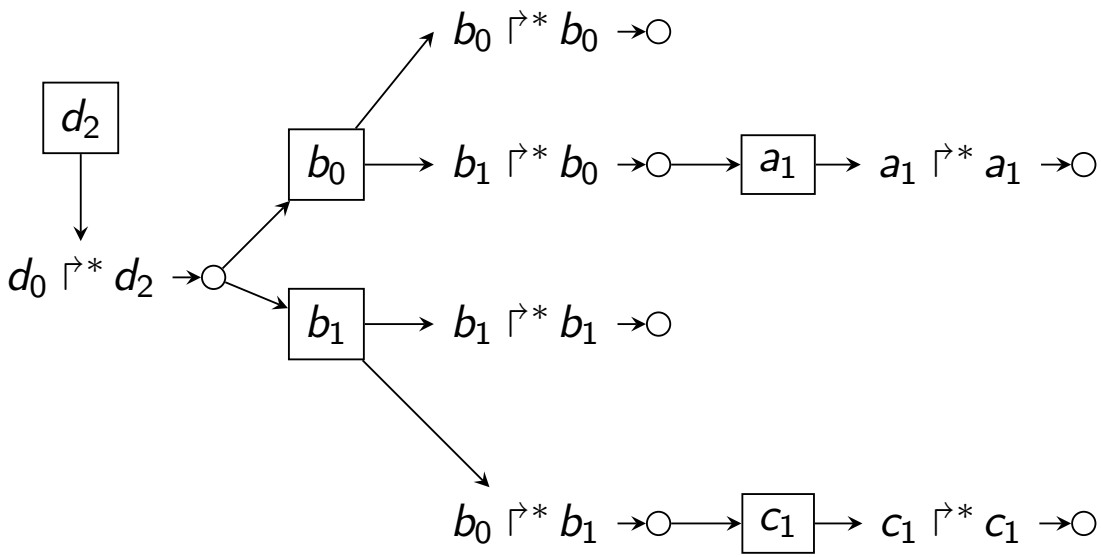


Under-approximation

Sufficient condition P:

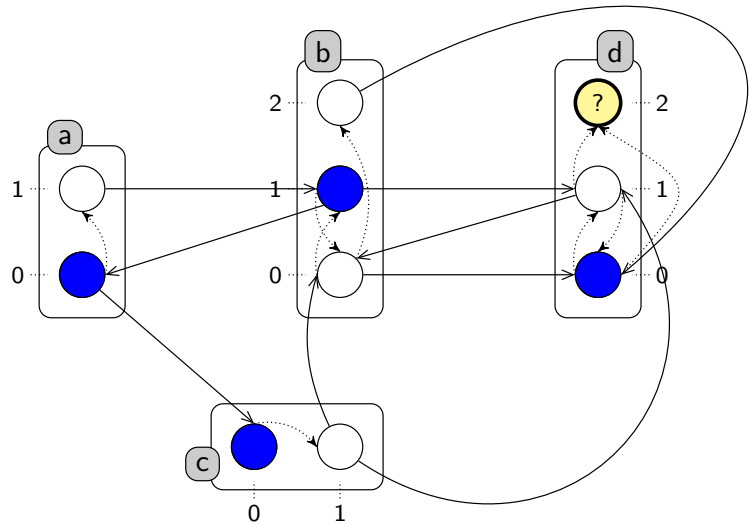
- no cycle
- each objective has a solution

P is true ⇒ R is true



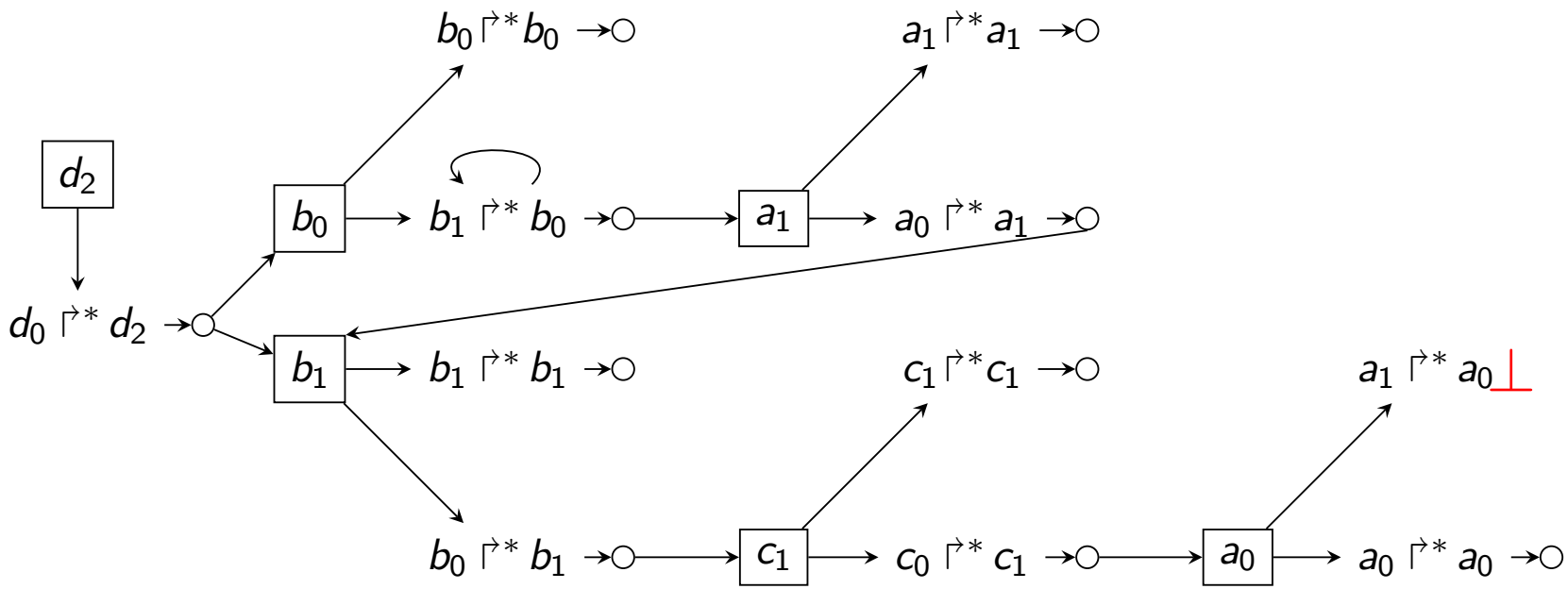
d_2	Required local state
$d_0 \lceil^* d_2$	Objective
○	Solution to an objective

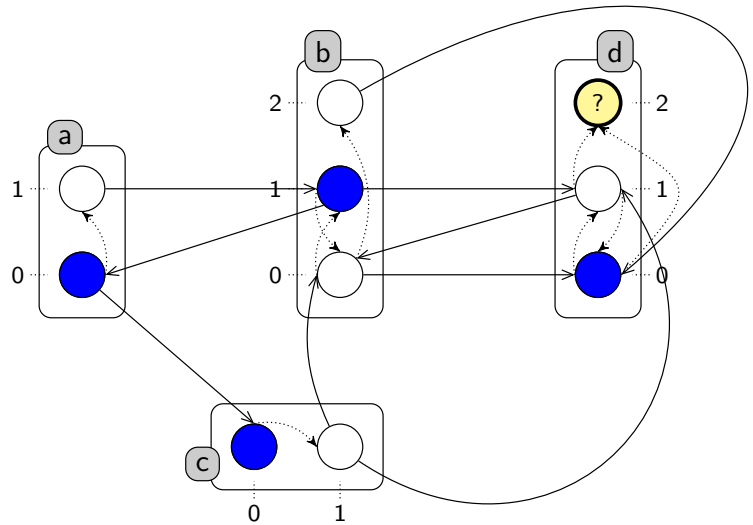
Under-approximation



Sufficient condition P:

- no cycle
- ~~each objective has a solution~~



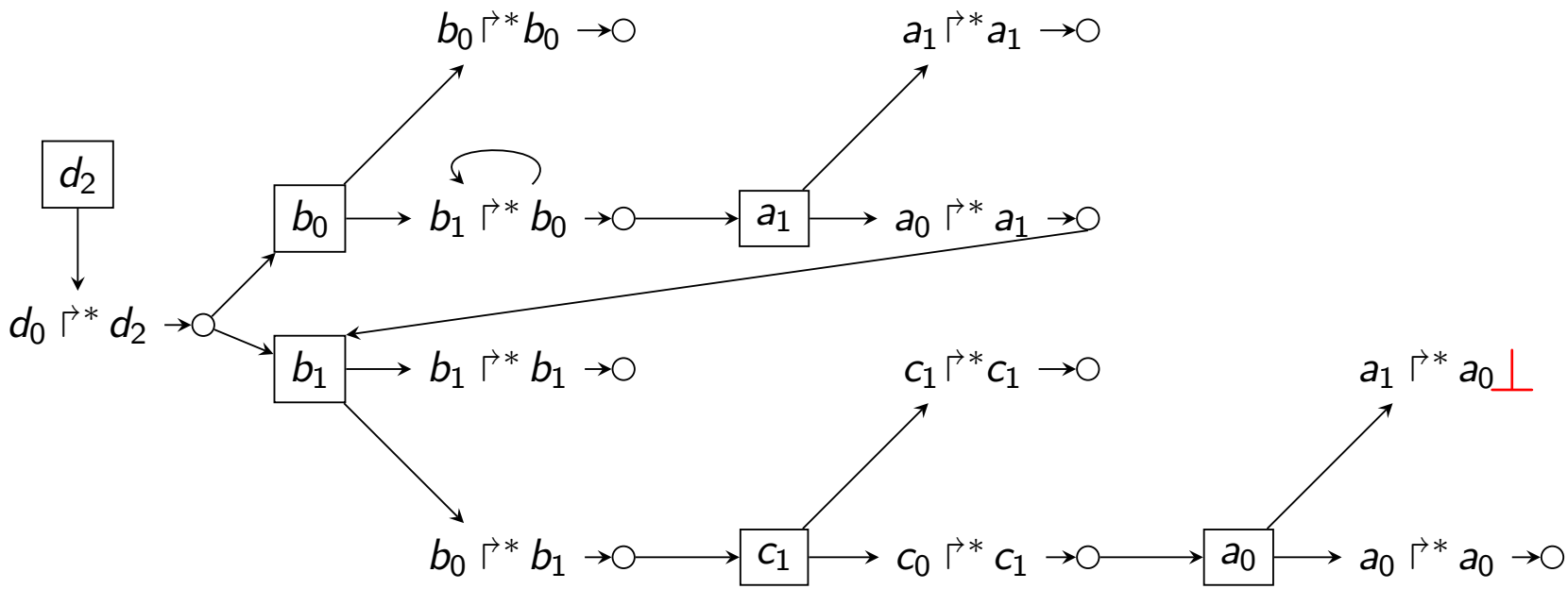


Under-approximation

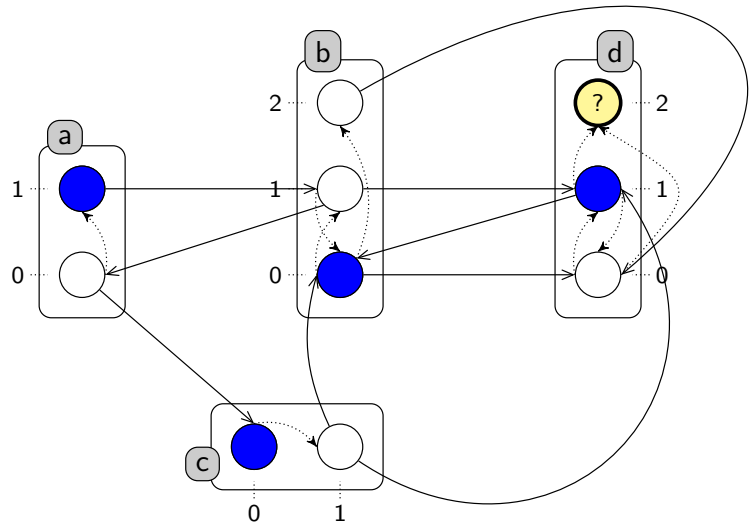
Sufficient condition P:

- no cycle
- ~~each objective has a solution~~

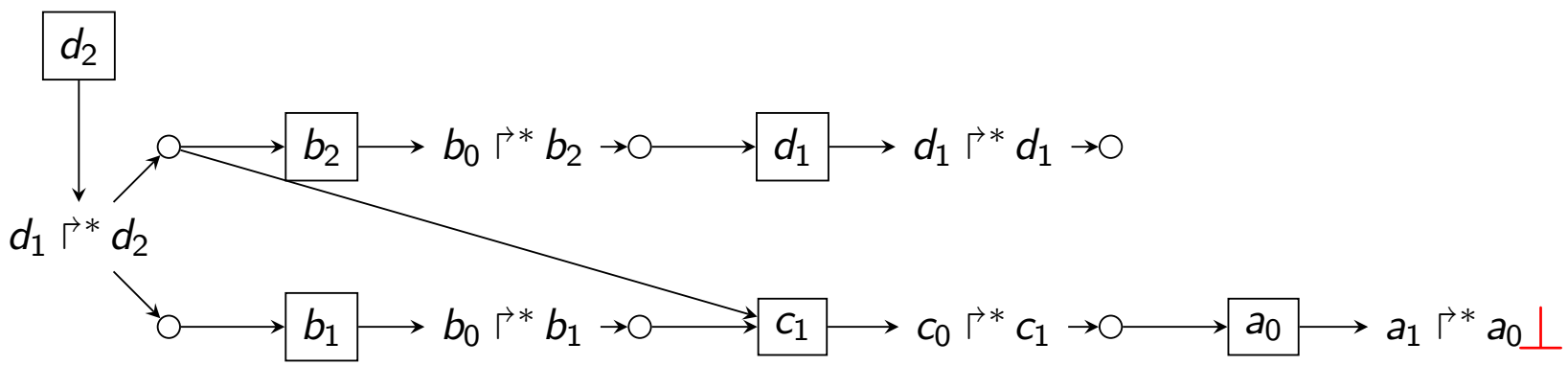
P is false ⇒ Inconclusive

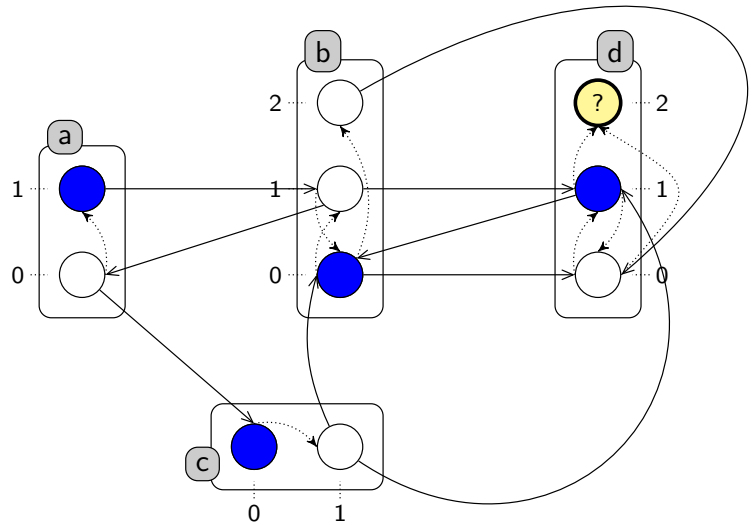


Over-approximation



Necessary condition Q:



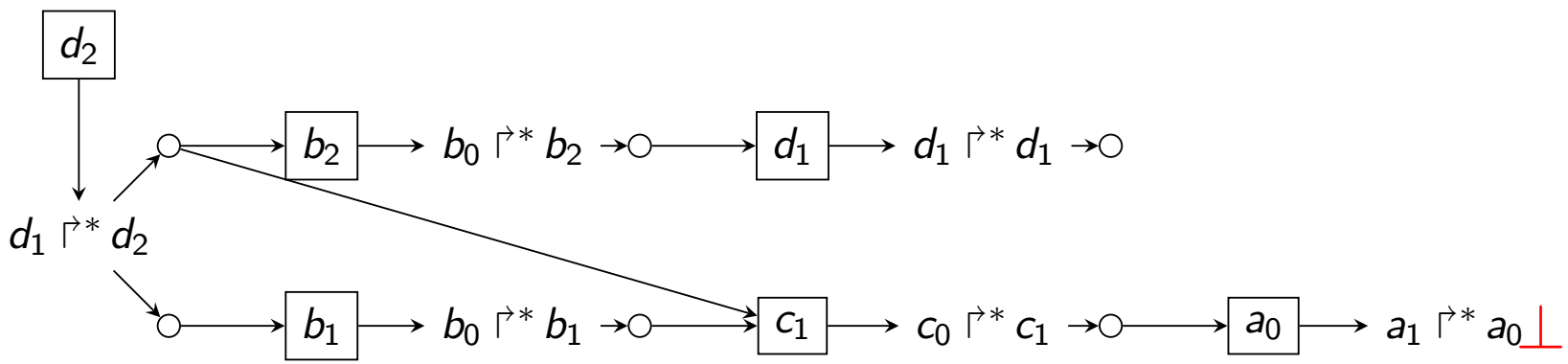


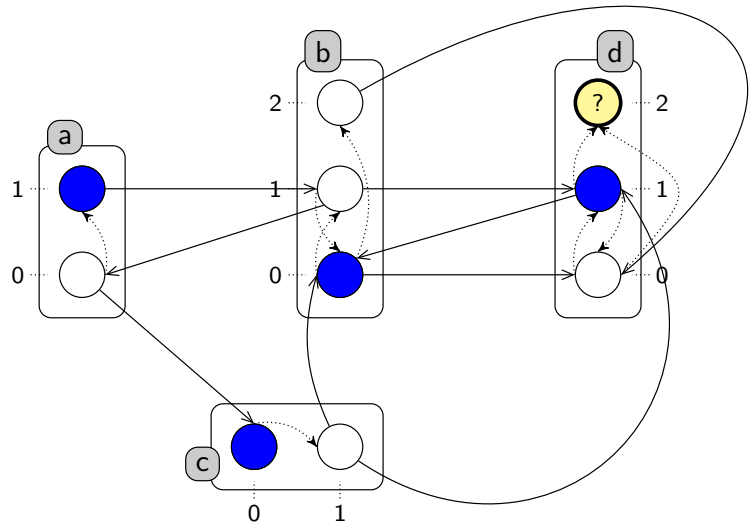
Over-approximation

Necessary condition Q:

There exists a traversal with no cycle

- objective → follow **one** solution
- solution → follow **all** local states
- local state → follow **all** objectives



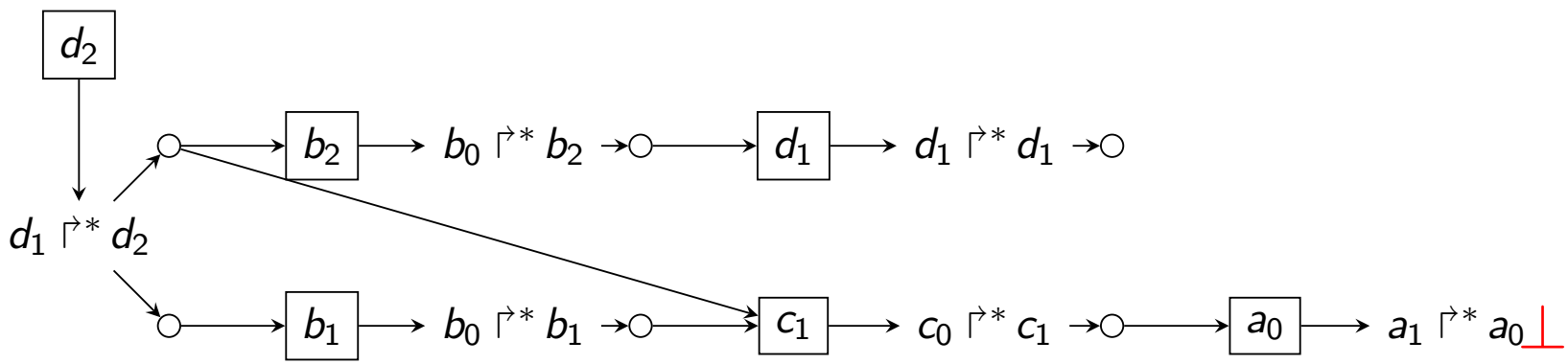


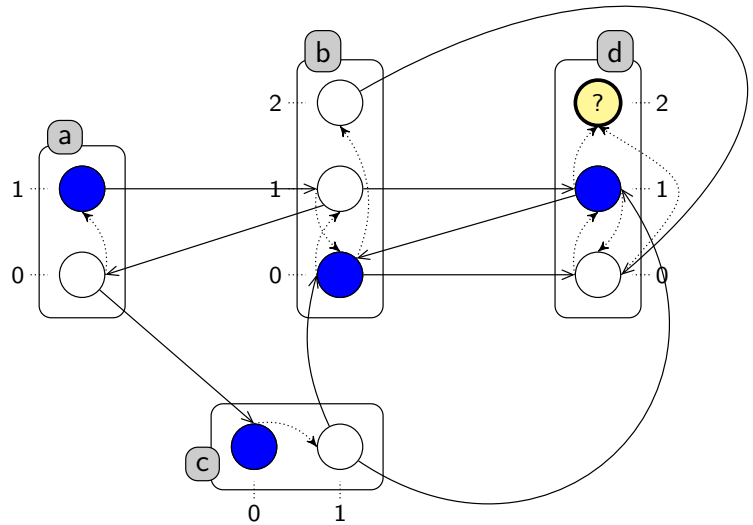
Over-approximation

Necessary condition Q:

~~There exists a traversal~~ with no cycle

- ~~objective → follow **one** solution~~
- solution → follow **all** local states
- local state → follow **all** objectives





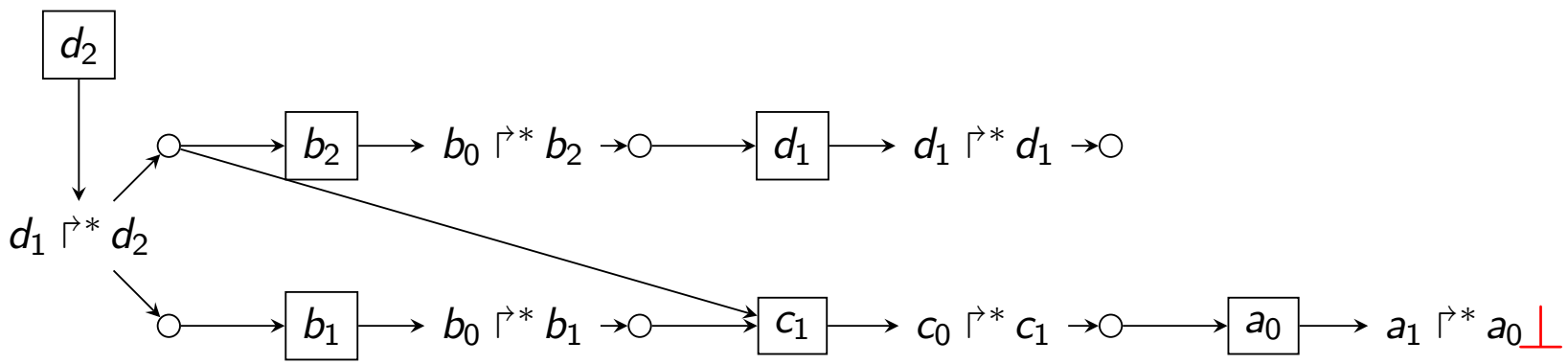
Over-approximation

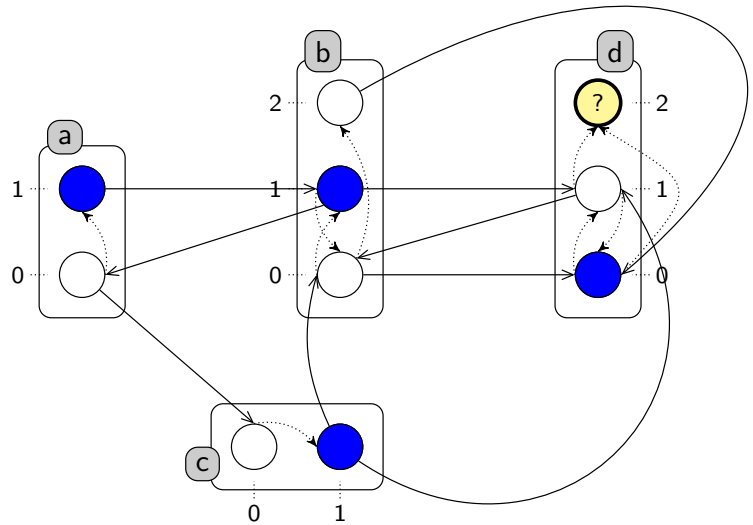
Necessary condition Q:

~~There exists a traversal~~ with no cycle

- ~~objective → follow **one** solution~~
- solution → follow **all** local states
- local state → follow **all** objectives

Q is false ⇒ R is false



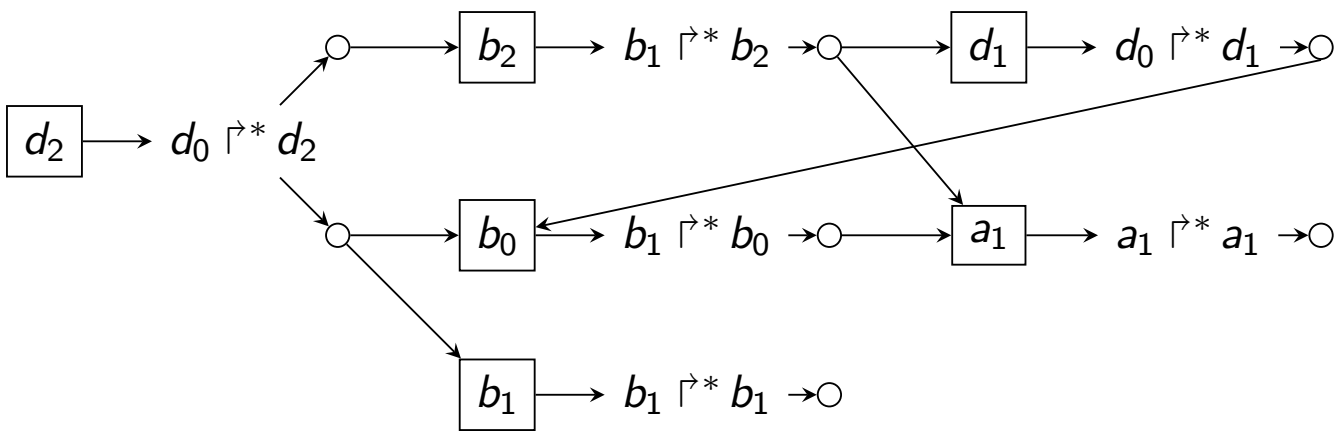


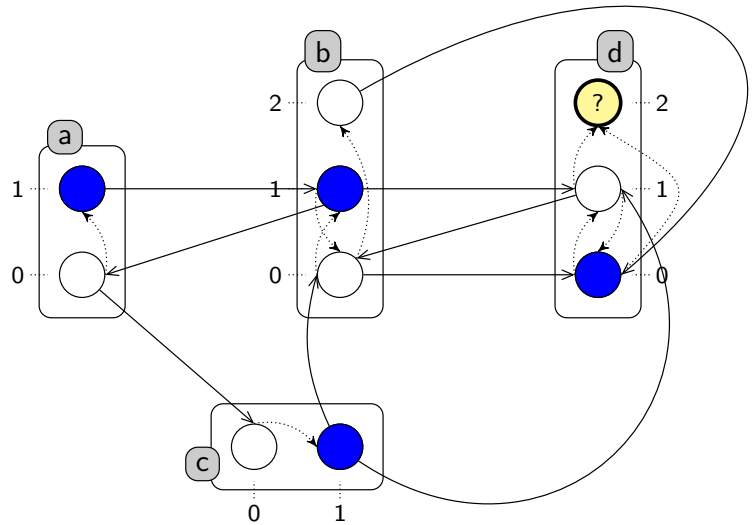
Over-approximation

Necessary condition Q:

There exists a traversal with no cycle

- objective → follow **one** solution
- solution → follow **all** local states
- local state → follow **all** objectives





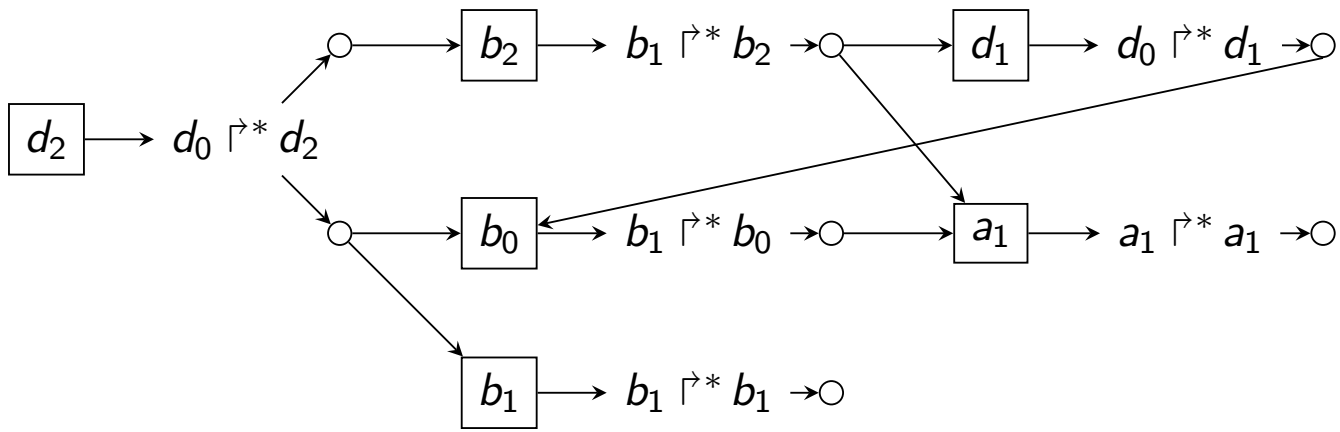
Over-approximation

Necessary condition Q:

There exists a traversal with no cycle

- objective → follow **one** solution
- solution → follow **all** local states
- local state → follow **all** objectives

Q is true ⇒ Inconclusive



Results

[Paulevé, *Computational Methods in Systems Biology*, 2017]
[Folschette *et al.*, *Theoretical Computer Science*, 2015]

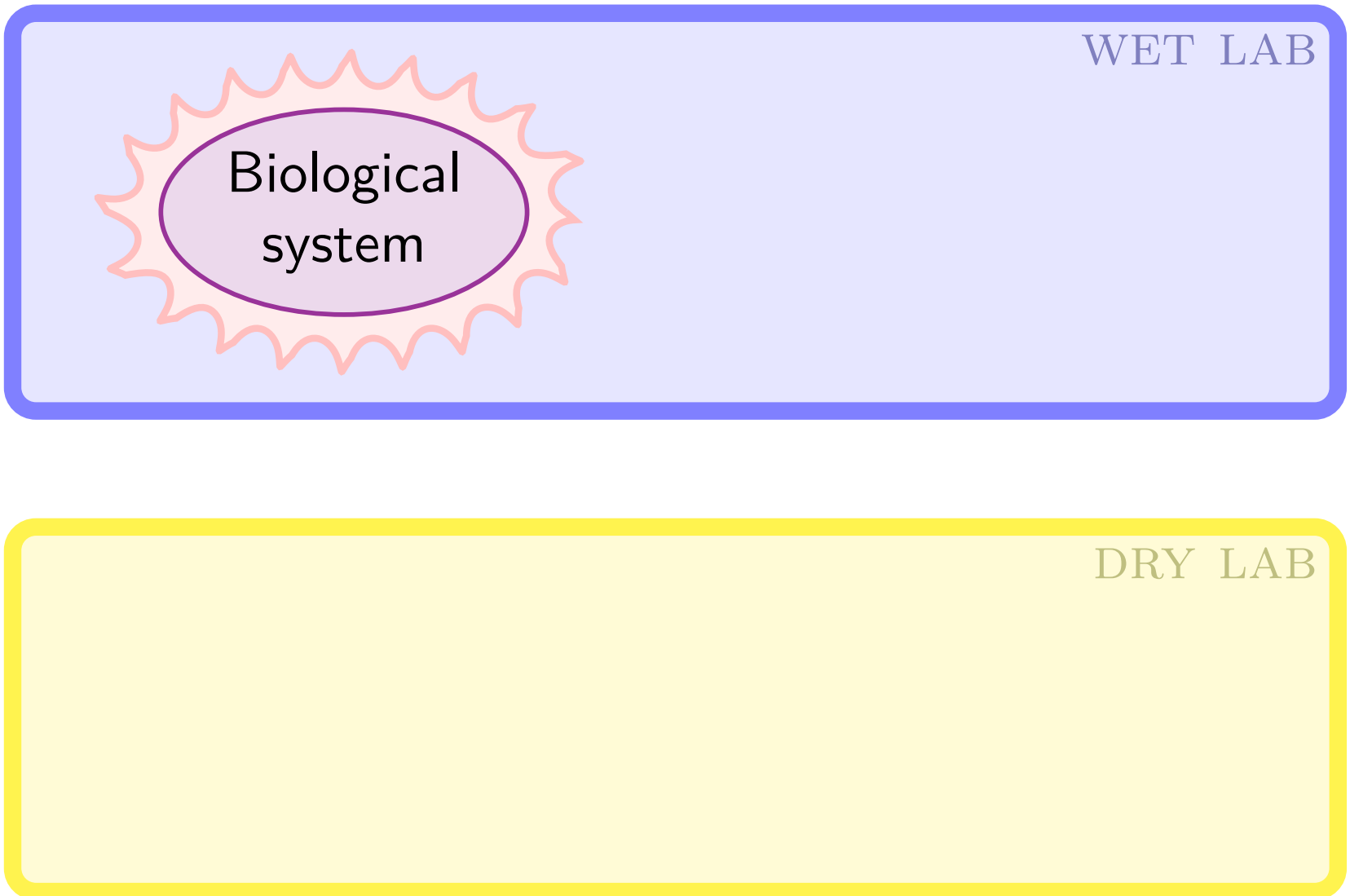
Model	nodes	NuSMV	ITS-reach	Pint
TCell-d	101	2s	0.5s	0.02s
		KO	16min	4.5s
RBE2F	370	KO	KO	0.2s
MAPK	309	KO	KO	48s

- KO = Out of memory
- 100% conclusiveness with Pint (no **Inconclusive**)

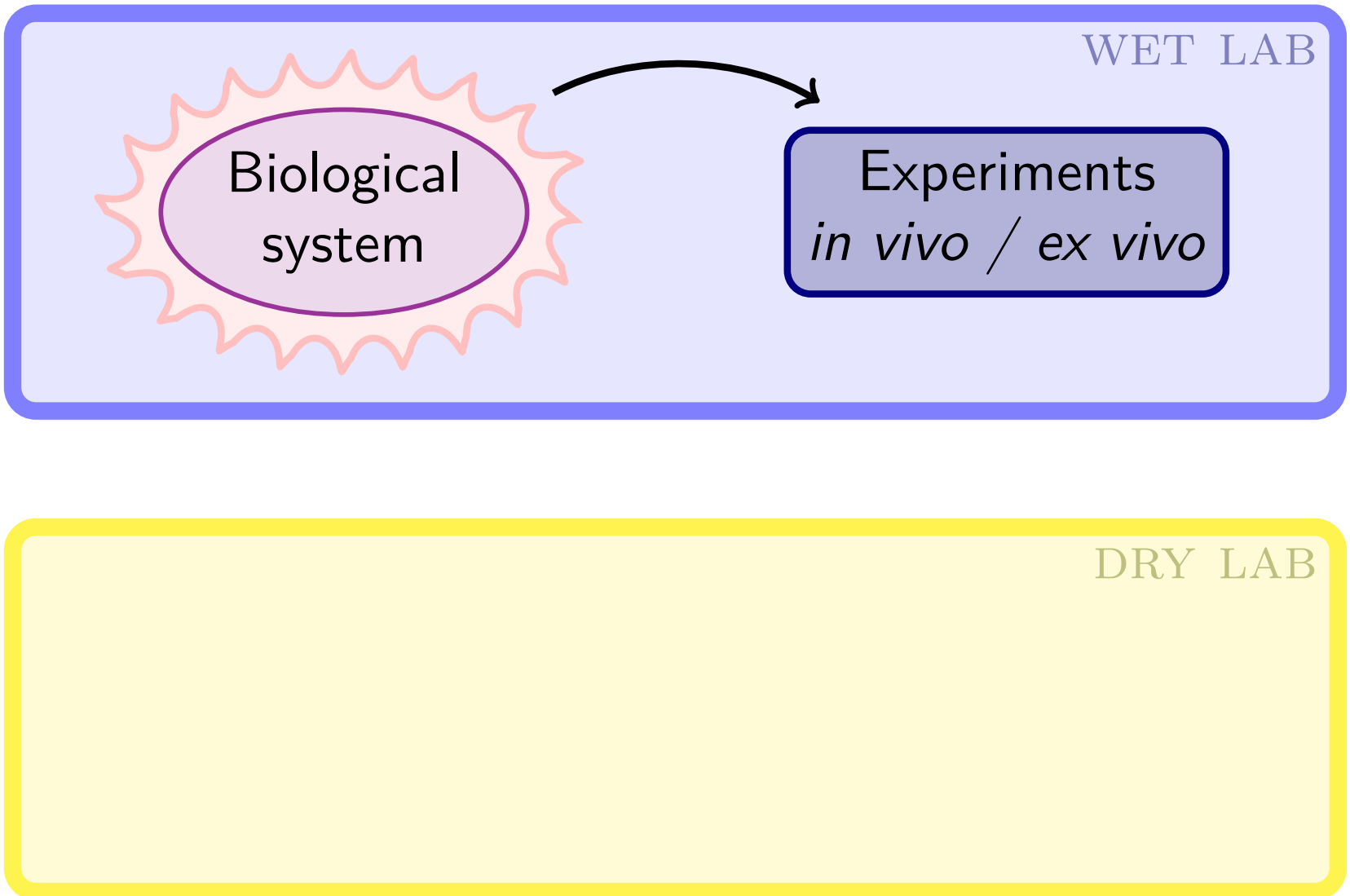
NuSMV: [Cimatti *et al.*, *CAV*, 2002. `http://nusmv.fbk.eu/`]
ITS-reach: [Thierry-Mieg, *TACAS*, 2015. `http://ddd.lip6.fr/itstools.php`]
Pint: [Paulevé, *Computational Methods in Systems Biology*, 2017. `http://loicpauleve.name/pint/`]

TCell-d: [Abou-Jaoudé *et al.*, *Frontiers in Bioengineering and Biotechnology*, 2015]
RBE2F: [Rougny, *et al.*, *BMC Systems Biology*, 2016]
MAPK: [Schoeberl *et al.*, *Nature biotechnology*, 2002]

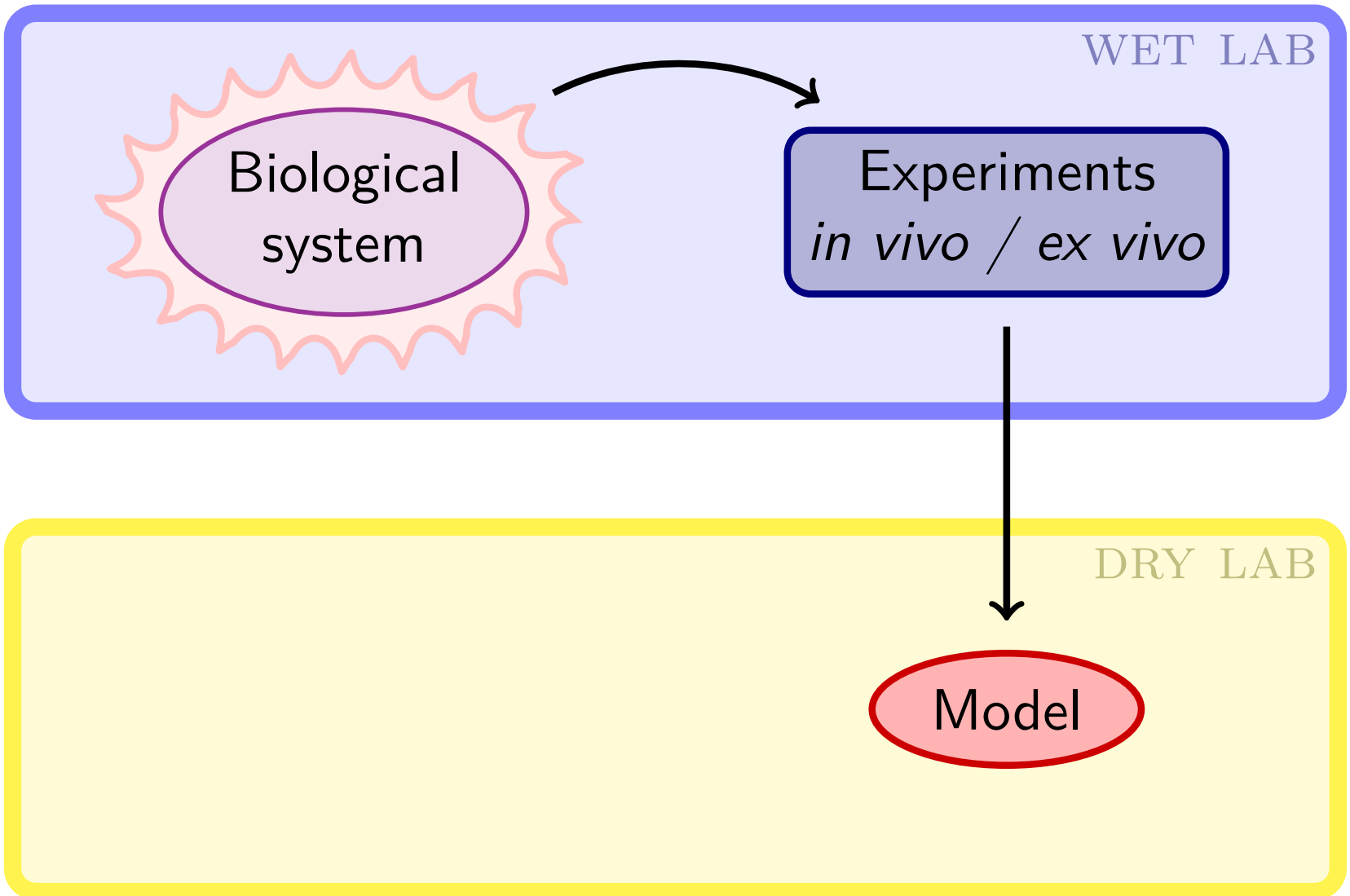
Experiments *in silico*



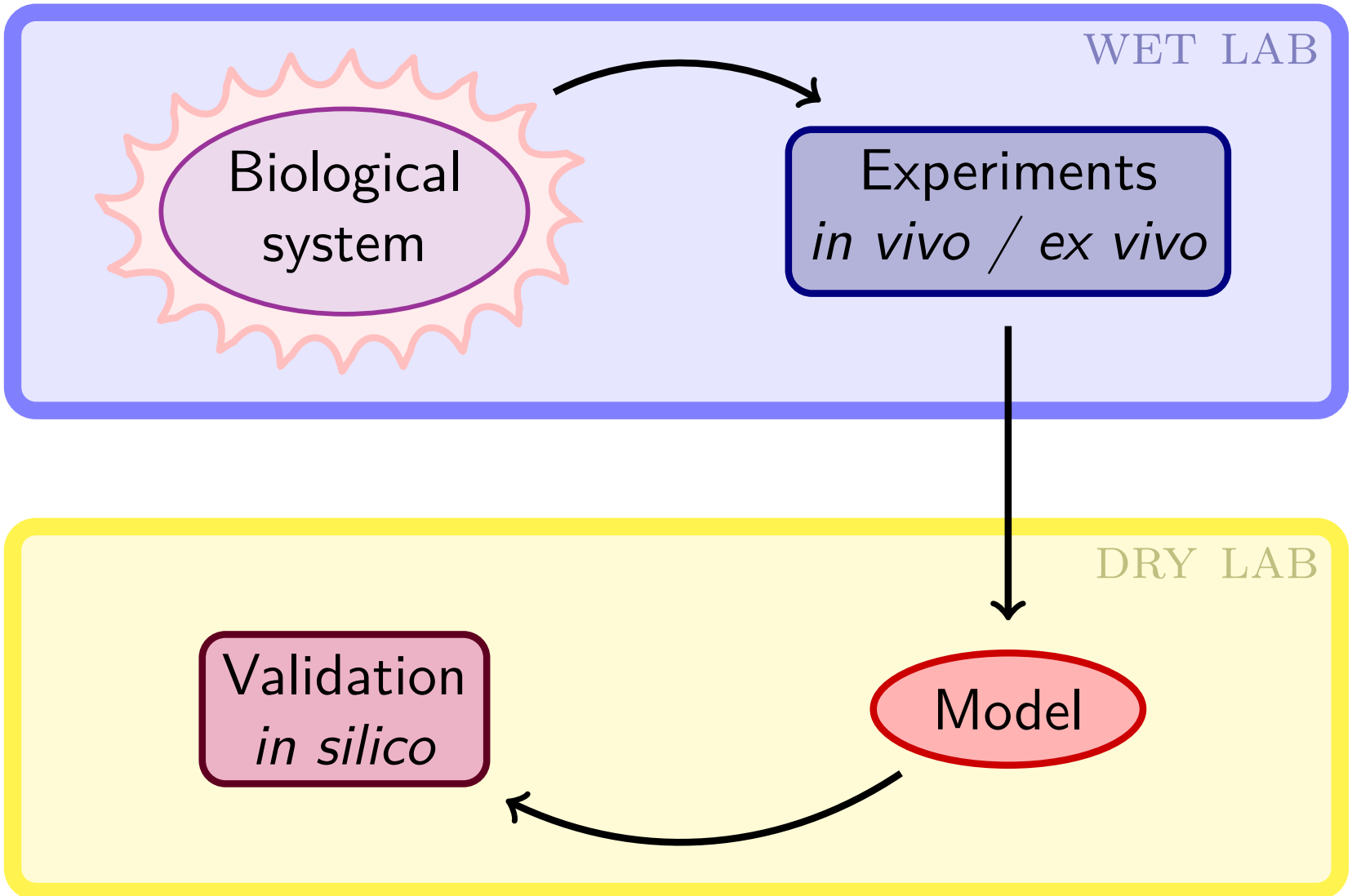
Experiments *in silico*



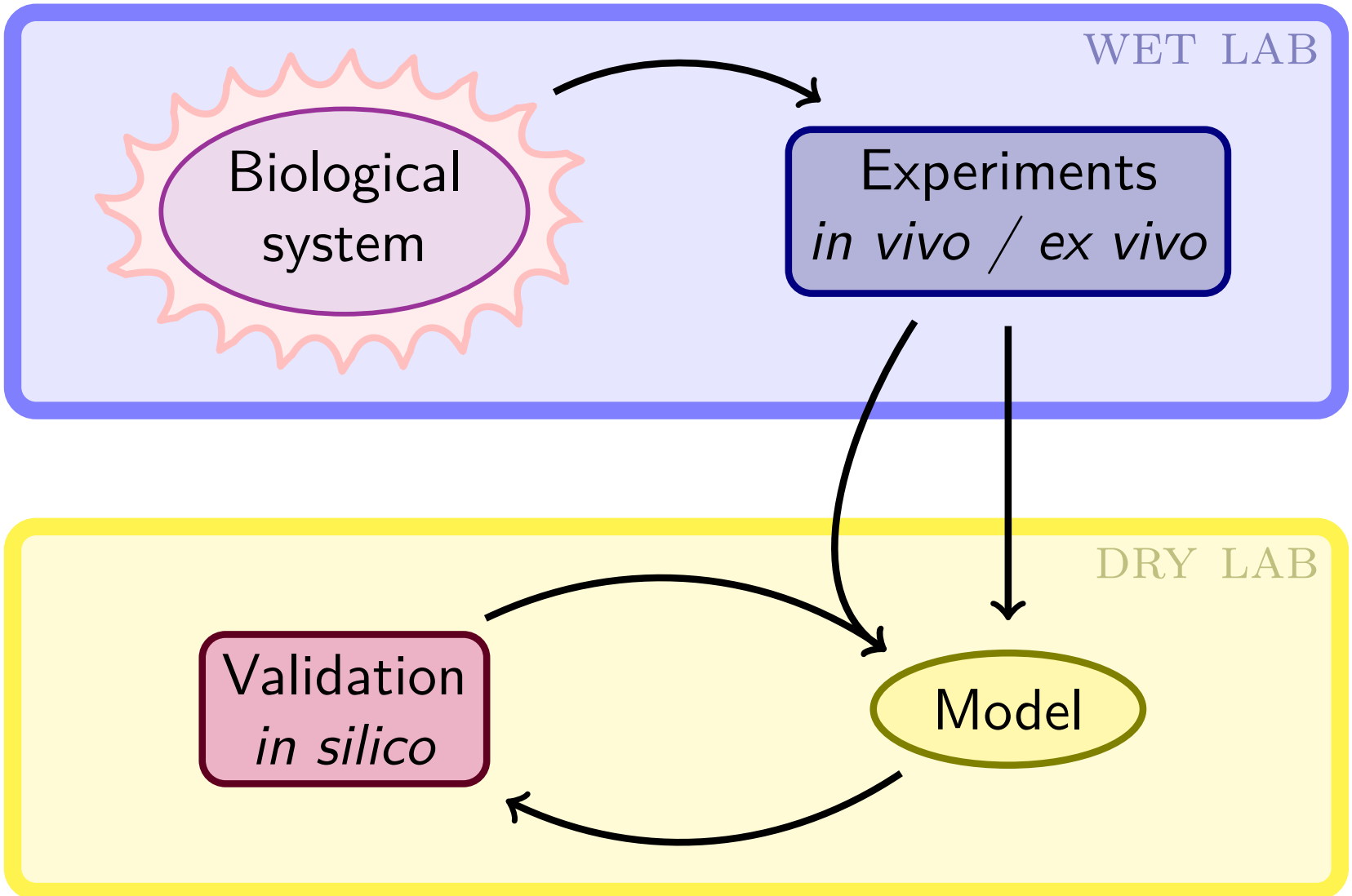
Experiments *in silico*



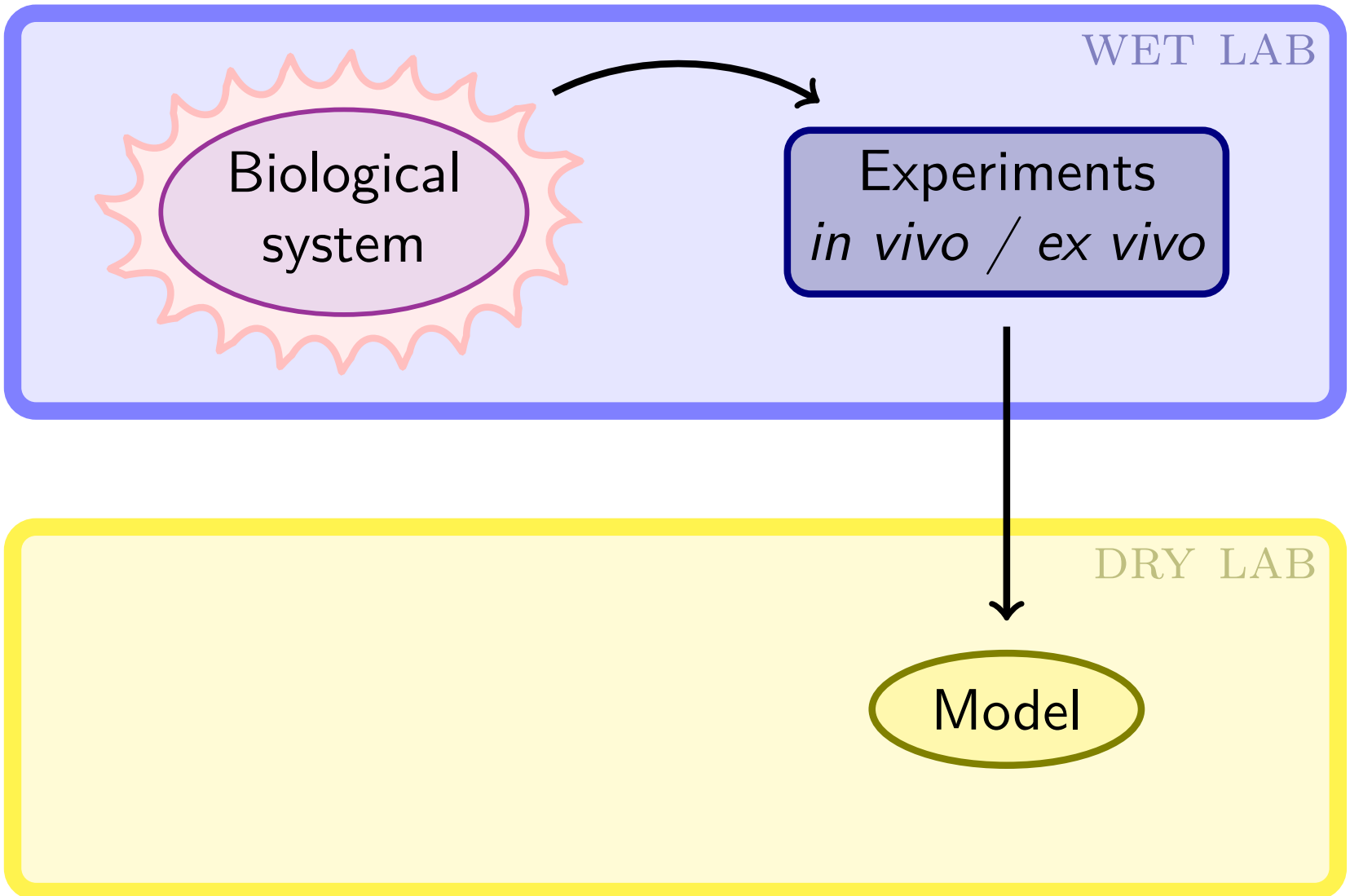
Experiments *in silico*



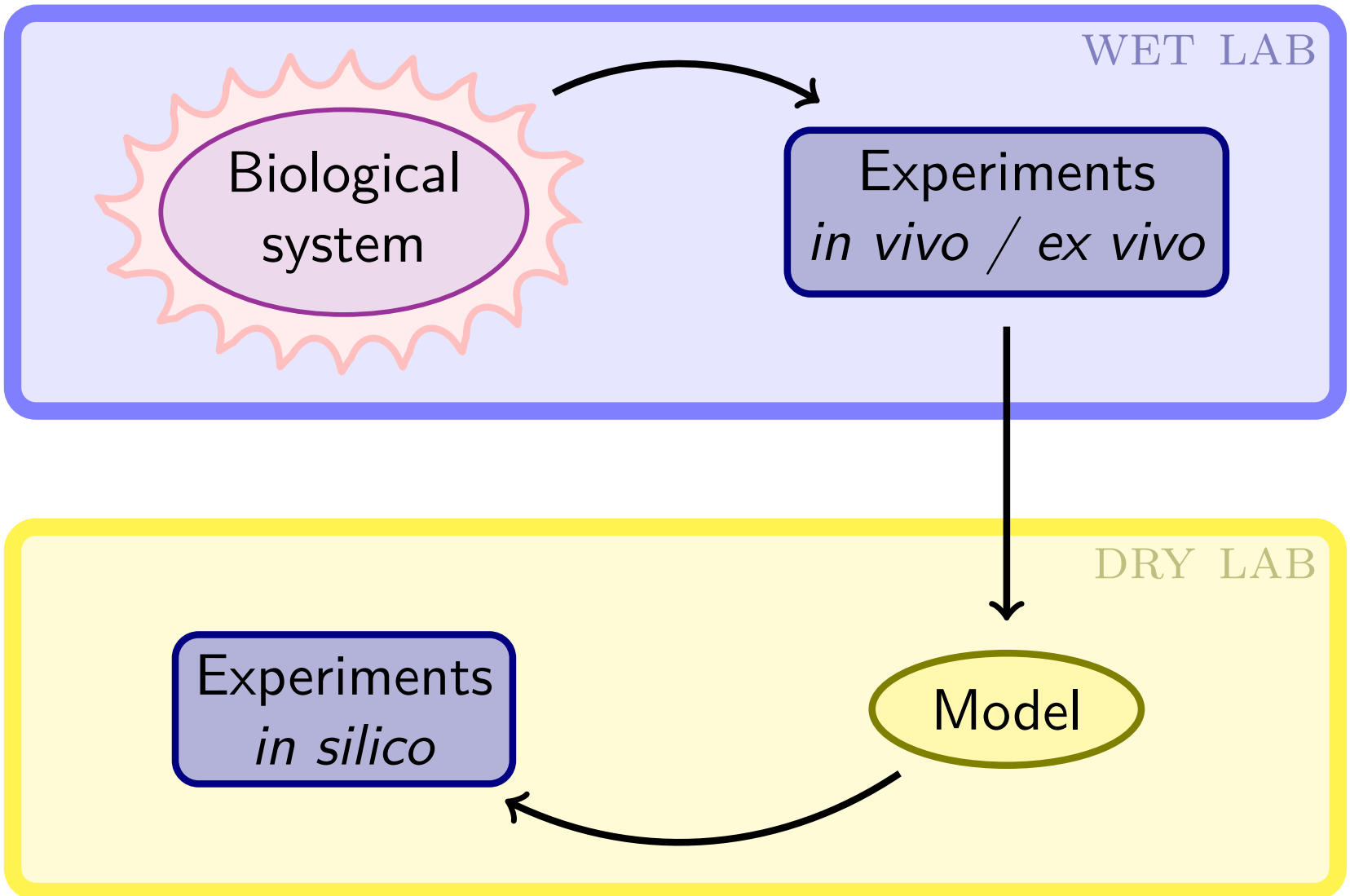
Experiments *in silico*



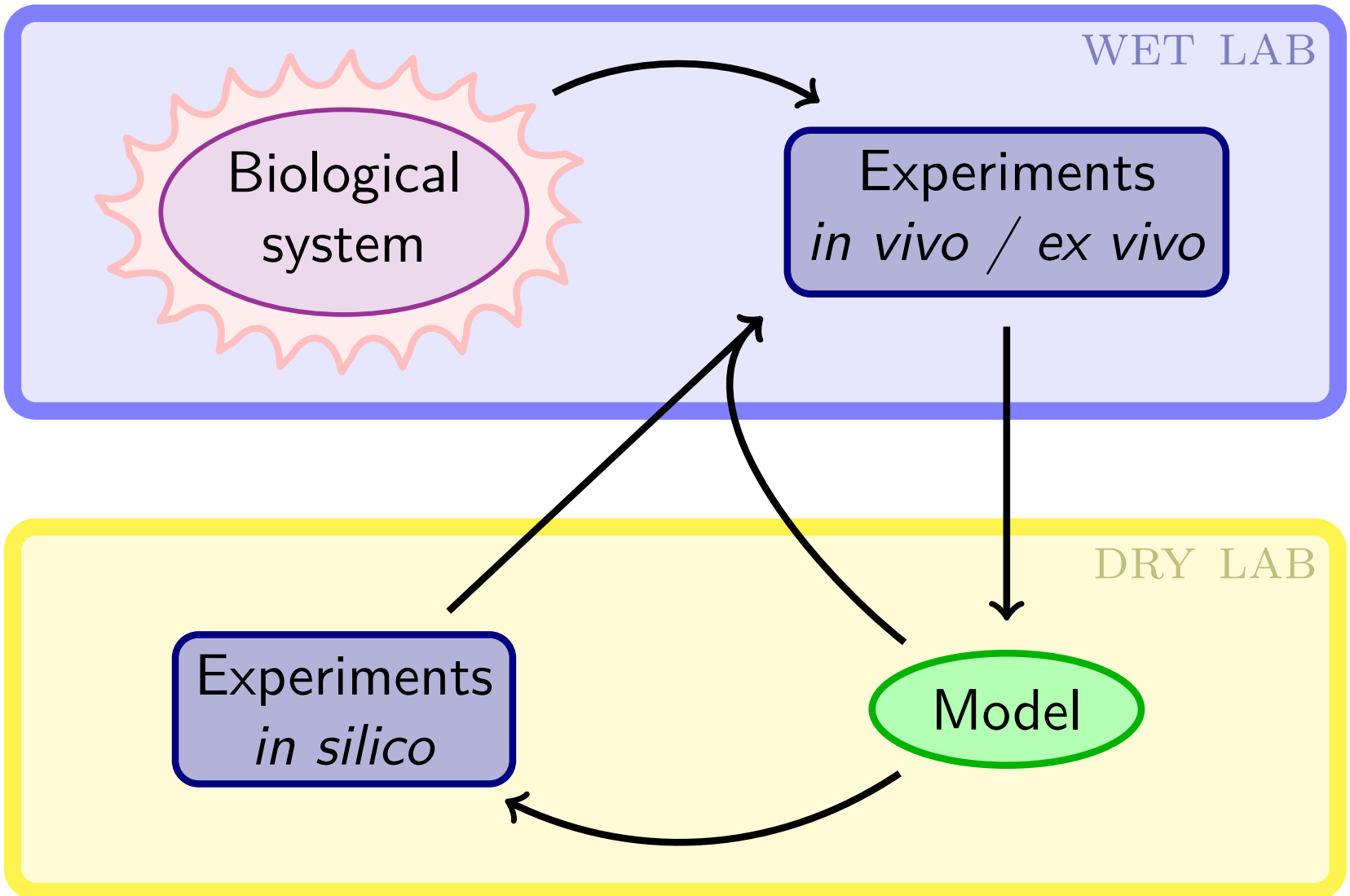
Experiments *in silico*



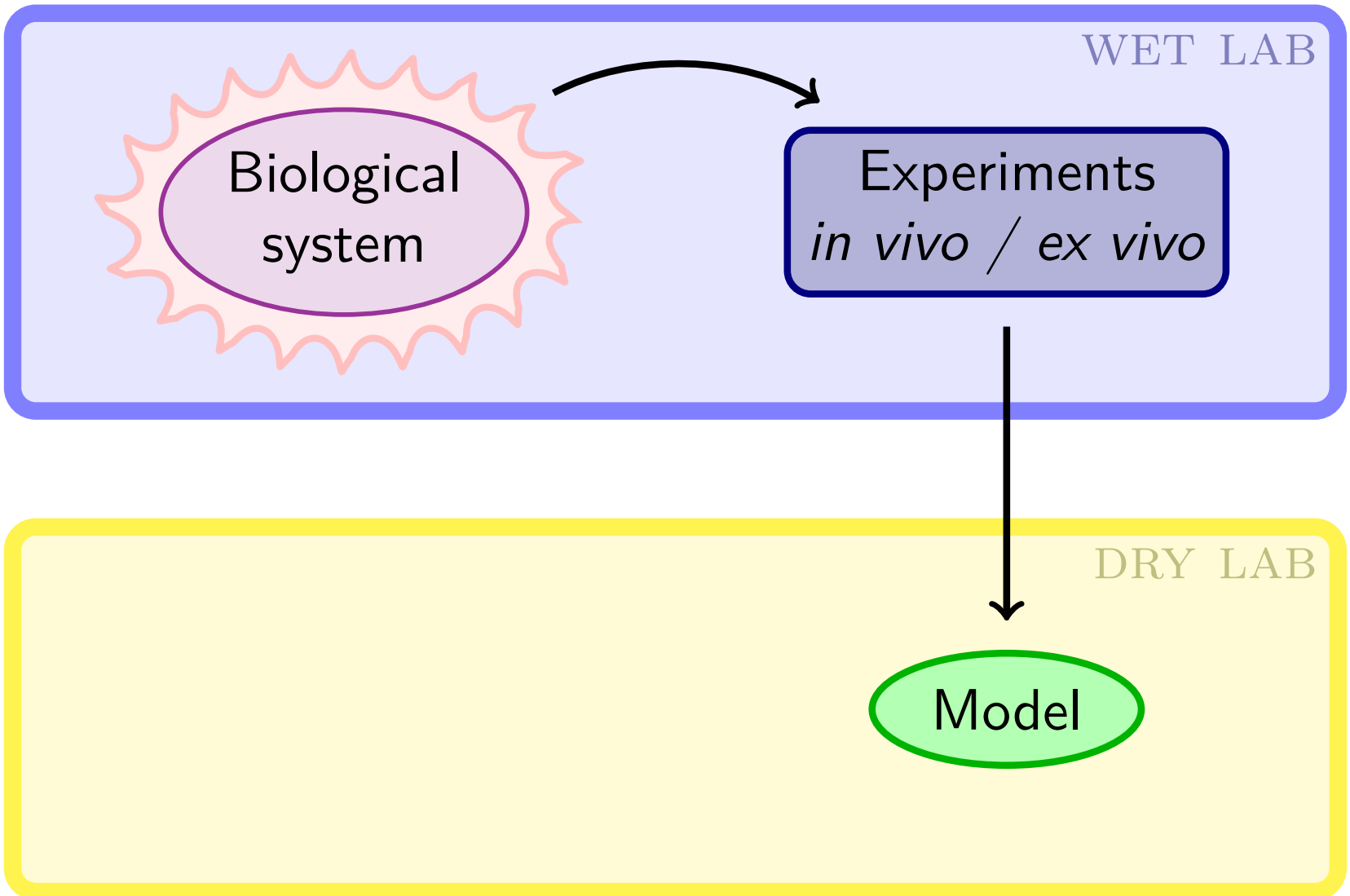
Experiments *in silico*



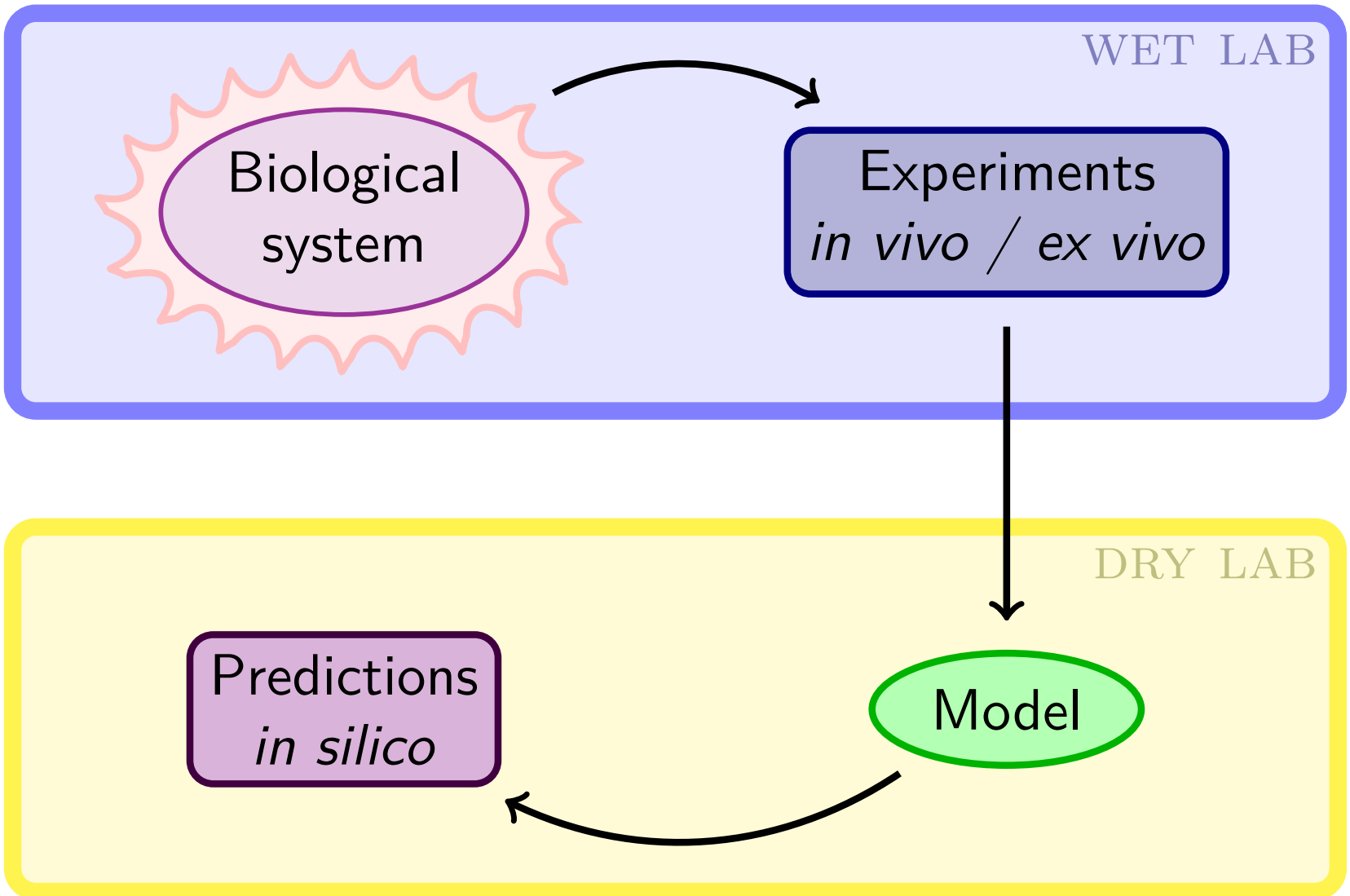
Experiments *in silico*



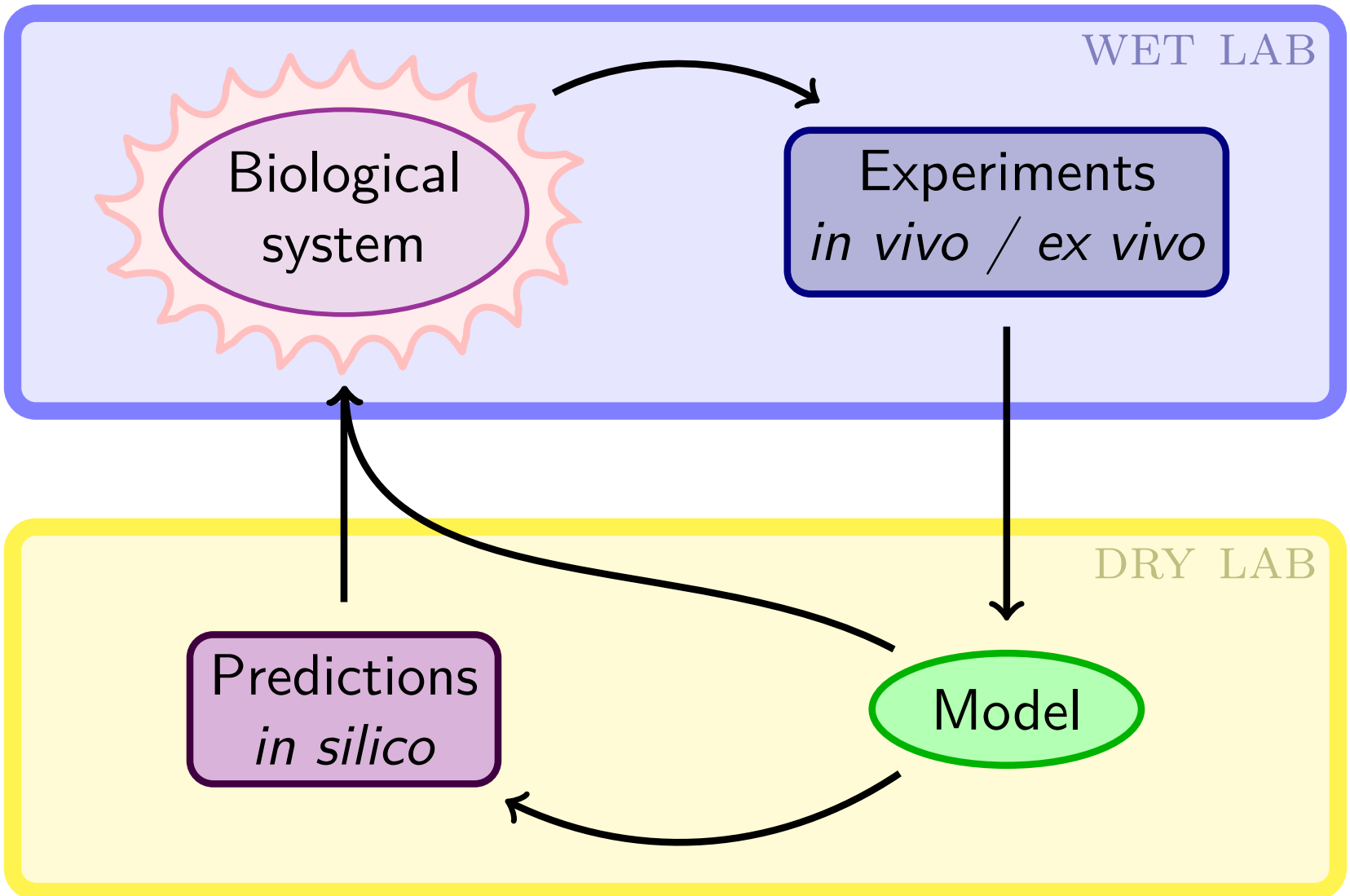
Experiments *in silico*



Experiments *in silico*

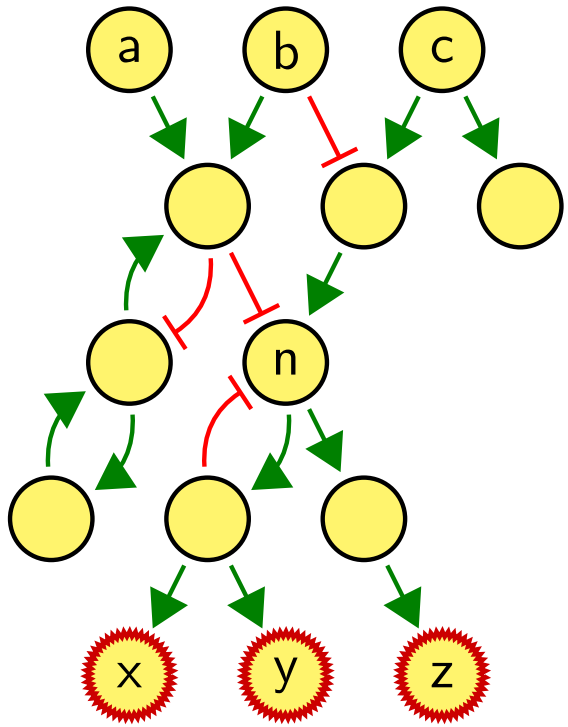


Experiments *in silico*



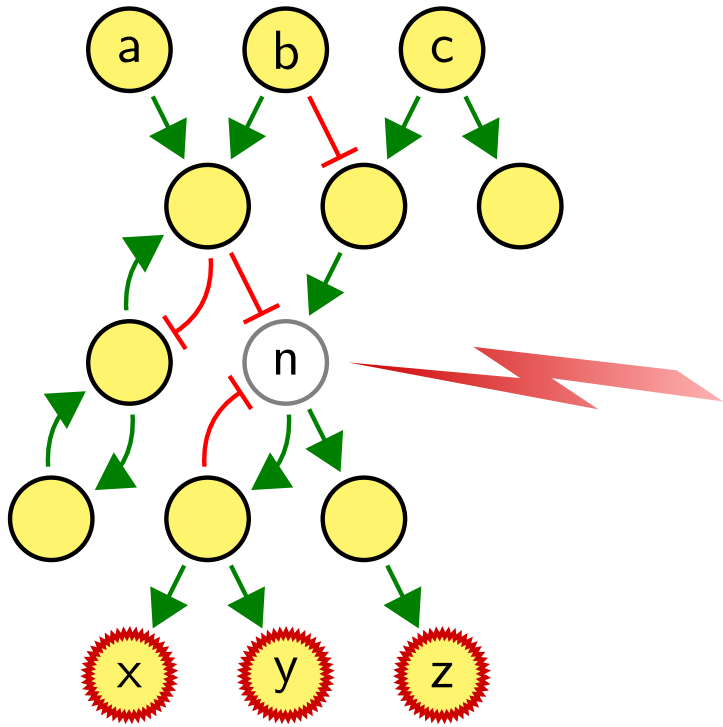
Gene Therapies

Find **therapeutic targets** in complex systems



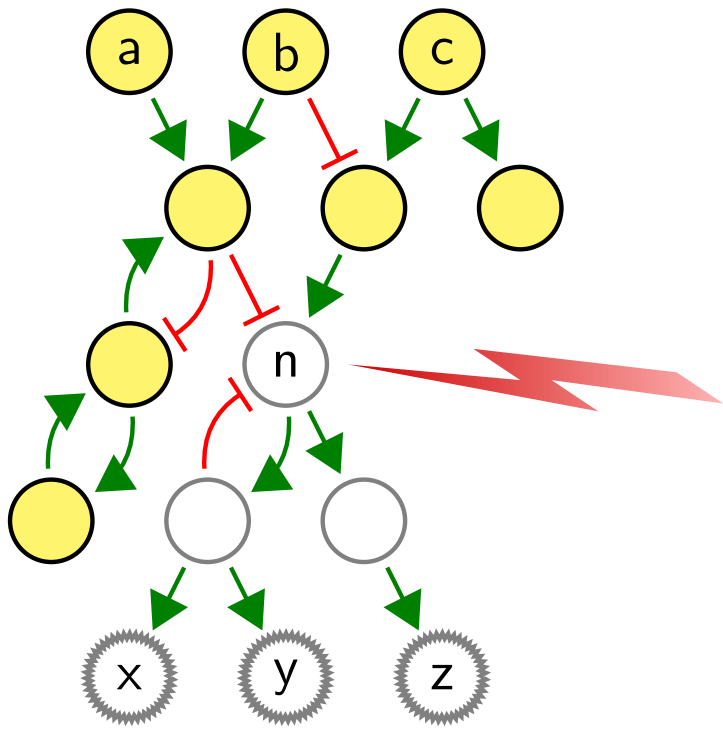
Gene Therapies

Find **therapeutic targets** in complex systems



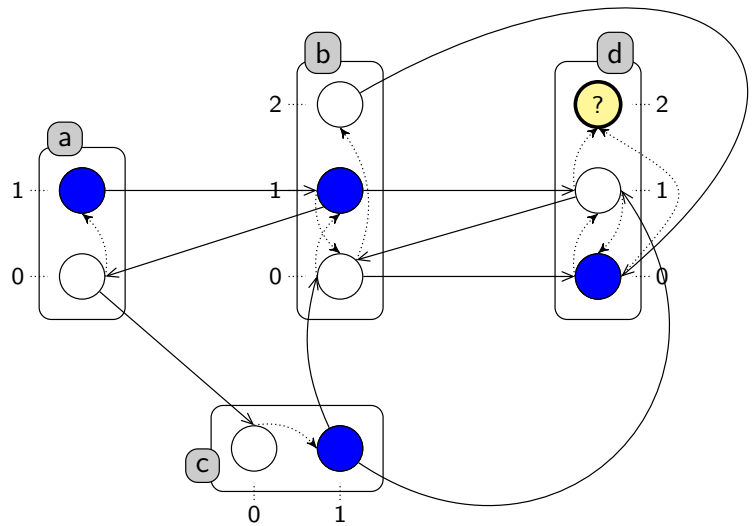
Gene Therapies

Find **therapeutic targets** in complex systems



Back to the Over-approximation

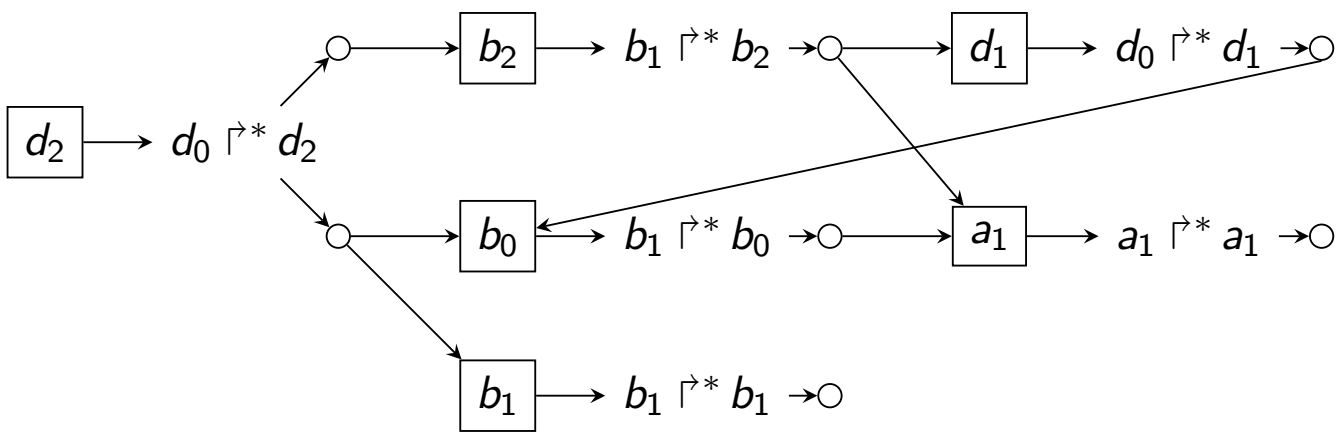
[Paulevé *et al.*, *Computer Aided Verification*, 2013]



Necessary condition Q:

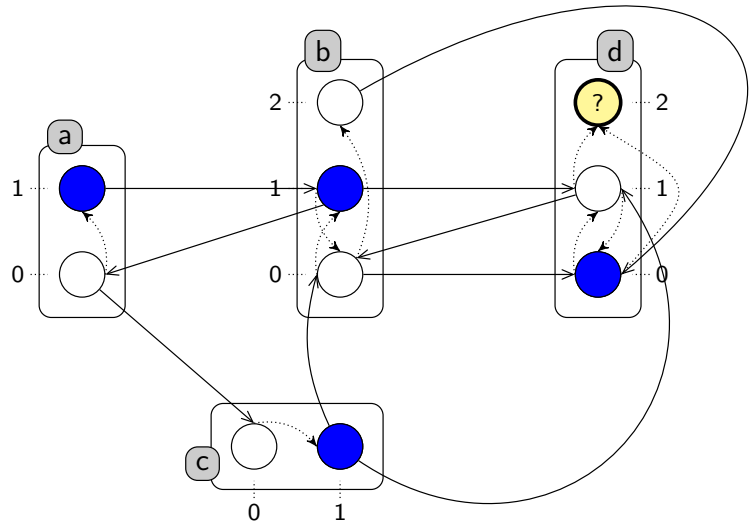
There exists a traversal with no cycle

- objective → follow **one** solution
- solution → follow **all** local states
- local state → follow **all** objectives



Back to the Over-approximation

[Paulevé *et al.*, *Computer Aided Verification*, 2013]

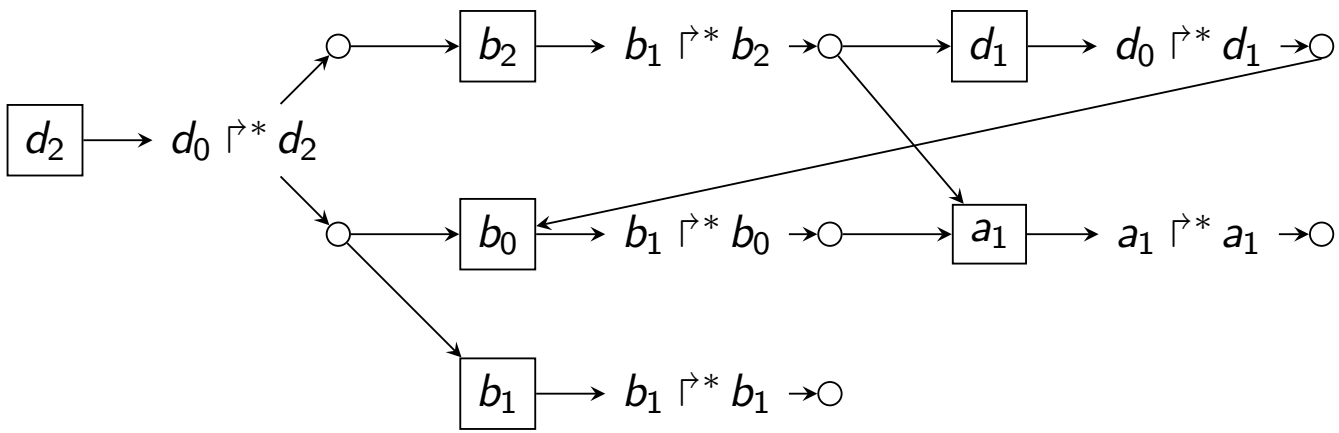


Necessary condition Q:

There exists a traversal with no cycle

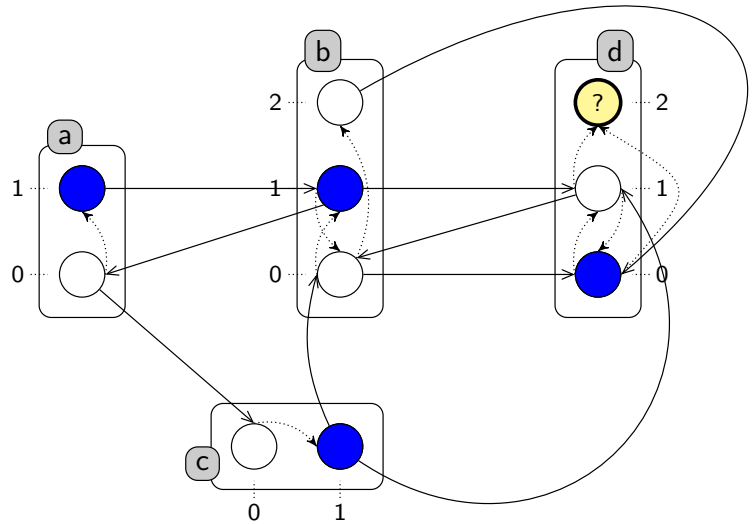
- objective → follow **one** solution
- solution → follow **all** local states
- local state → follow **all** objectives

Q is true ⇒ Inconclusive



Back to the Over-approximation

[Paulevé *et al.*, *Computer Aided Verification*, 2013]

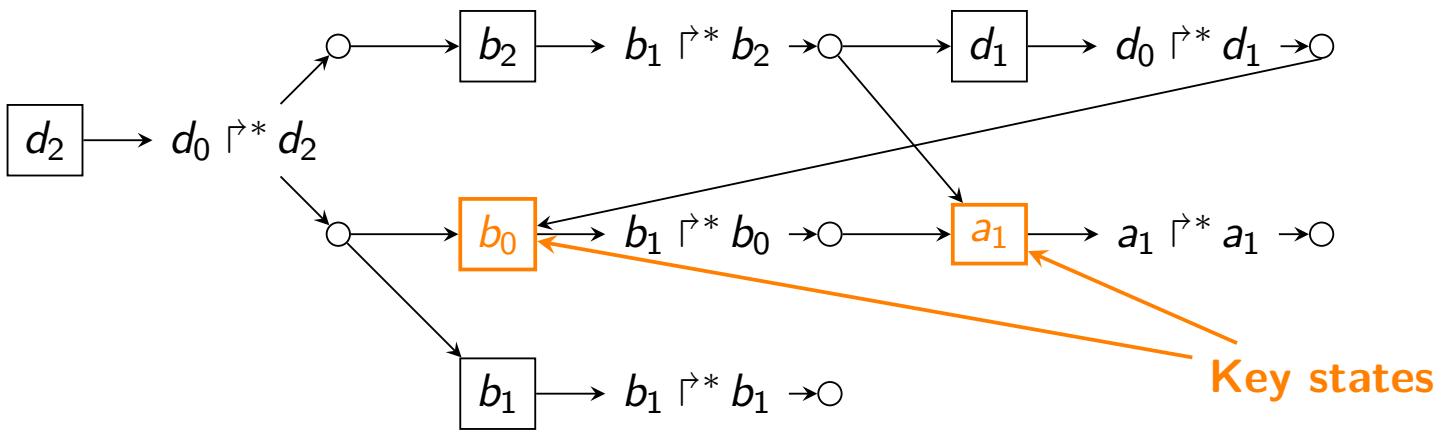


Necessary condition Q:

There exists a traversal with no cycle

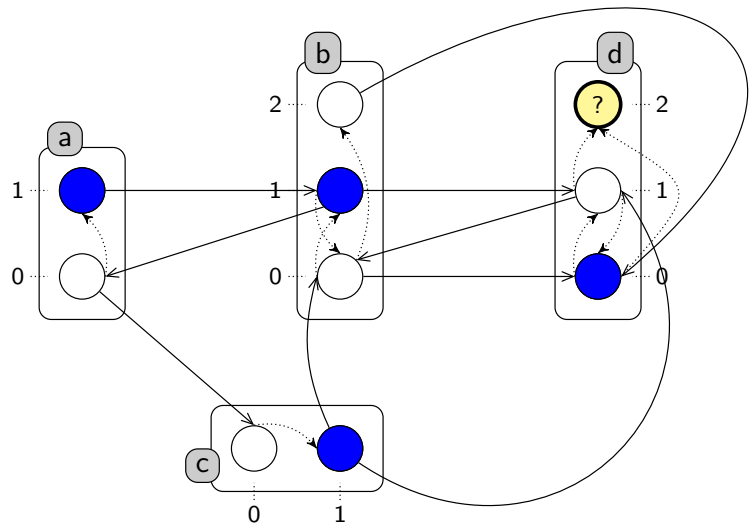
- objective → follow **one** solution
- solution → follow **all** local states
- local state → follow **all** objectives

Q is true ⇒ Inconclusive



Back to the Over-approximation

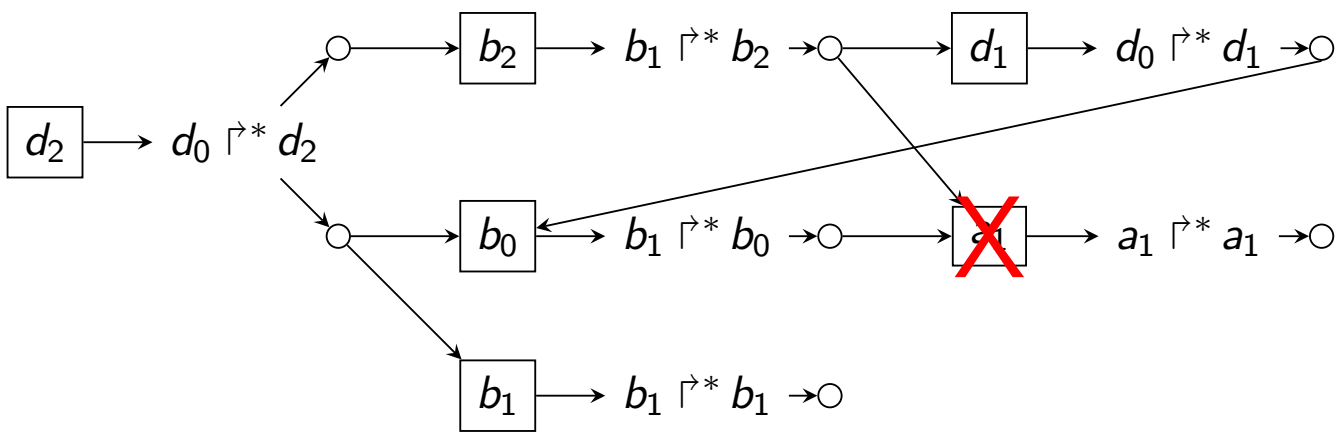
[Paulevé *et al.*, *Computer Aided Verification*, 2013]



Necessary condition Q:

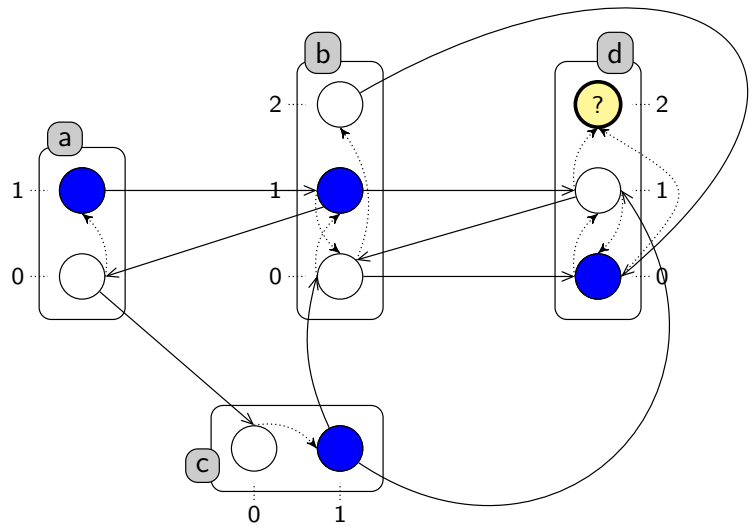
There exists a traversal with no cycle

- objective → follow **one** solution
- solution → follow **all** local states
- local state → follow **all** objectives



Back to the Over-approximation

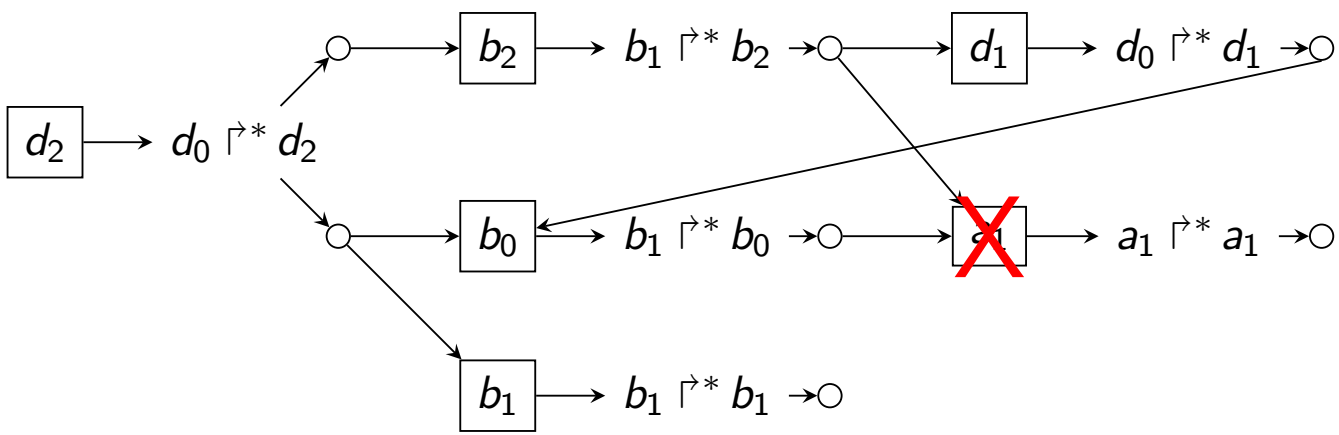
[Paulevé *et al.*, *Computer Aided Verification*, 2013]



Necessary condition Q:

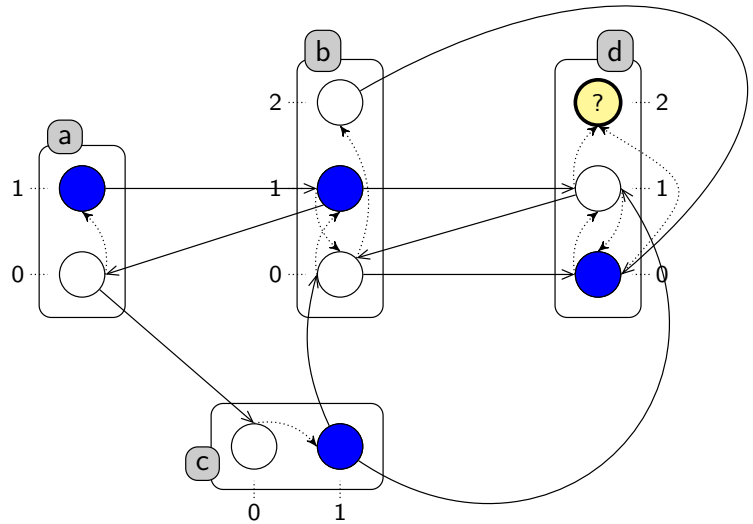
~~There exists a traversal~~ with no cycle

- objective → follow **one** solution
- ~~solution → follow **all** local states~~
- local state → follow **all** objectives



Back to the Over-approximation

[Paulevé *et al.*, *Computer Aided Verification*, 2013]

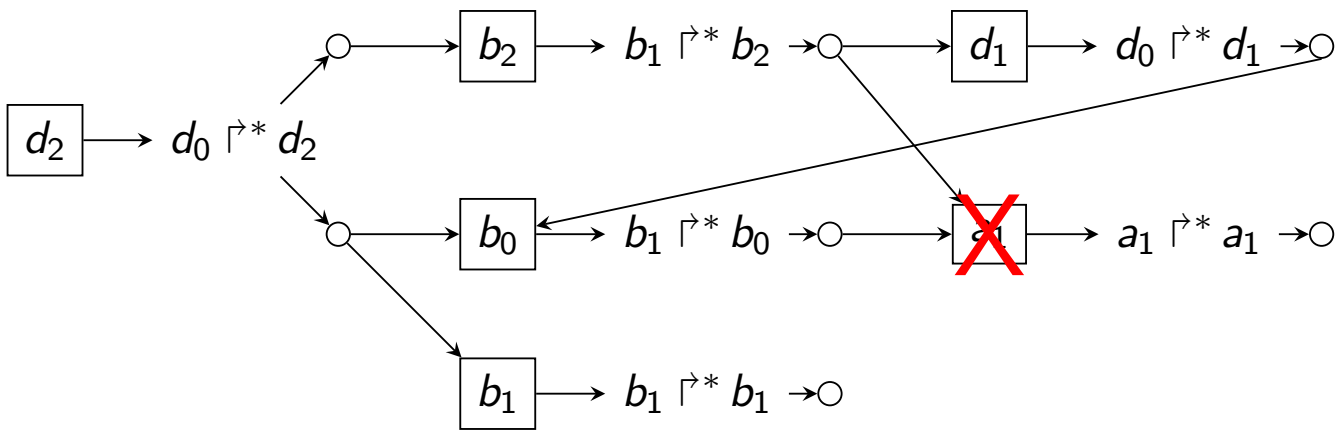


Necessary condition Q:

~~There exists a traversal~~ with no cycle

- objective → follow **one** solution
- ~~solution → follow **all** local states~~
- local state → follow **all** objectives

Q is false ⇒ R is false



Summary & Conclusion

Modeling of René Thomas for **Regulatory networks**

$\Rightarrow$ How to study its dynamics?

Dynamic analysis

- Classical approach: state-graph computation
- Precise but computational (exponential)

  $\Rightarrow$ Large models are intractable

Static analysis

- Much faster (polynomial) but lesser range of results
- Reachability analysis:

  $\Rightarrow$ Outperforms classical approaches

  $\Rightarrow$ Very few inconclusive cases

Bibliography

- René Thomas. Boolean formalization of genetic control circuits. *Journal of Theoretical Biology*, Volume 42, n°3, 563–85, 1973.
- Hidde de Jong. Modeling and simulation of genetic regulatory systems: a literature review. *Journal of Computational biology* 9(1), 67–103. 2002.
- René Thomas. On the Relation Between the Logical Structure of Systems and Their Ability to Generate Multiple Steady States or Sustained Oscillations. In *Numerical Methods in the Study of Critical Phenomena* Volume 9, Jean Della Dora, Jacques Demongeot, Bernard Lacolle, editors, 180–93. Springer Berlin Heidelberg, 1981.
- Loïc Paulevé, Morgan Magnin, Olivier Roux. Static analysis of biological regulatory networks dynamics using abstract interpretation. *Mathematical Structures in Computer Science*. 2012.
- Maxime Folschette, Loïc Paulevé, Morgan Magnin, Olivier Roux. Sufficient Conditions for Reachability in Automata Networks with Priorities. *Theoretical Computer Science*, special edition *From Computer Science to Biology and Back*, Volume 608, Part 1, 66–83. Elsevier, 2015.

Thank you