

**Constraint-Based Methods for Bioinformatics Workshop**

# **Identification of Bifurcations in Biological Regulatory Networks using Answer-Set Programming**

Louis Fippo Fitime<sup>1</sup>

`louis.fippo-fitime@irccyn.ec-nantes.fr`

**Joint work with:**

Olivier Roux<sup>1</sup>, Carito Guziolowski<sup>1</sup> and Loïc Paulevé<sup>2</sup>

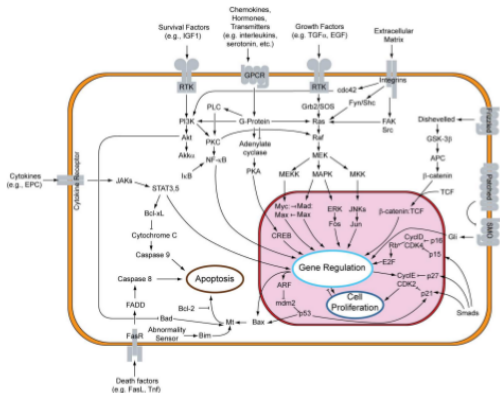
<sup>1</sup> MeForBio / IRCCyN / École Centrale de Nantes (Nantes, France)

<sup>2</sup> Bioinfo / LRI / CNRS (Paris, France)

`olivier.roux@irccyn.ec-nantes.fr`

`carito.guziolowski@irccyn.ec-nantes.fr`

`loic.pauleve@lri.fr`

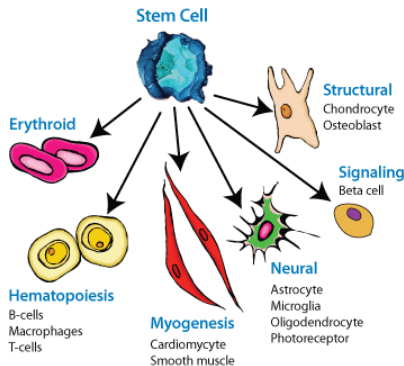


[Liu et al, in *Journal of Bioinformatics and Computational Biology*, 2012]

- Cellular processes are driven by networks of biological interactions.
- Formal modelling and analysis of Biological Regulatory Network.
- Static analysis of properties.

# Motivation

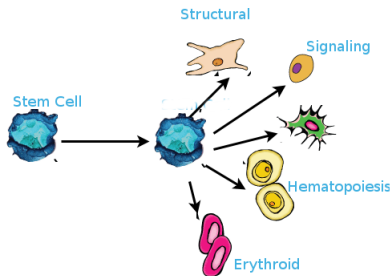
stem cell differentiation



[<https://www.systembio.com/stem-cell-research/differentiation-reporters/overview>]

- Loss of capability **differentiation**
- Which transitions (operations) are responsible of **Bifurcations** ?
- From which state ?

## Contribution



- Bifurcation transitions can be expressed in temporal logic formula

- **"Relaxation"** :

verification of CTL formula (**PSPACE (complete)**)

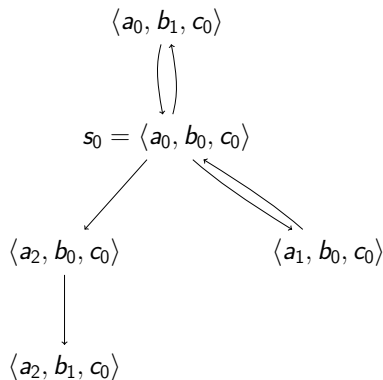
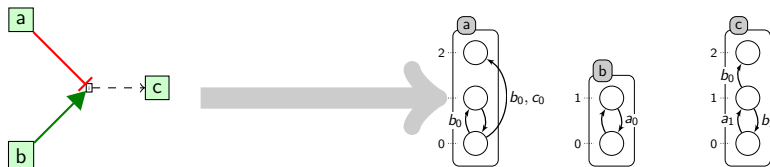
⇒ NP problem which can be easily expressed in SAT/ASP

- **Contribution** :

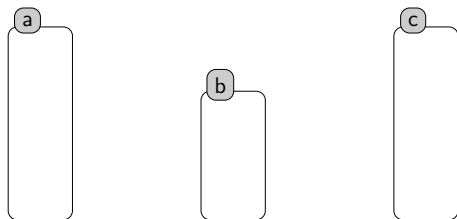
$\left\{ \begin{array}{l} \text{AI (Artificial intelligence)} \\ + \\ \text{AI (Abstract interpretation)} \end{array} \right.$

⇒ **formal approximation of bifurcations**

## Biological networks modelling

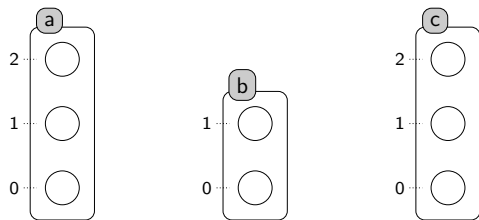


## Automata Networks



Automata: components *a, b, c*

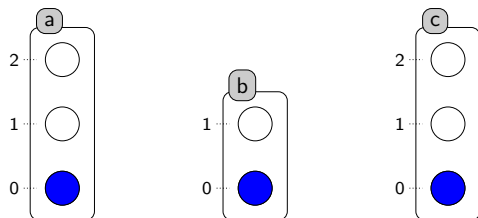
## Automata Networks



**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

## Automata Networks



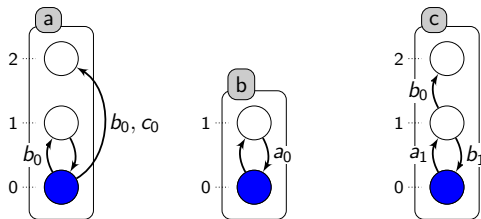
**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

**States:** sets of active local states  $\langle a_0, b_0, c_0 \rangle$



## Automata Networks



$$s_0 = \langle a_0, b_0, c_0 \rangle$$

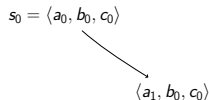
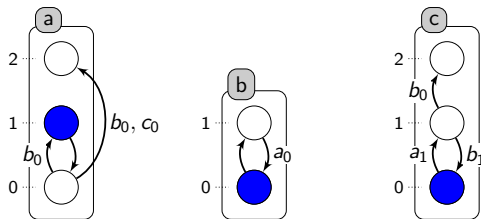
**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

**States:** sets of active local states  $\langle a_0, b_0, c_0 \rangle$

**Transitions:** dynamics  $t_1 = a_0 \xrightarrow{b_0} a_1, t_2 = a_1 \rightarrow a_0, t_3 = a_0 \xrightarrow{b_0, c_0} a_2, t_4 = b_0 \rightarrow b_1$

## Automata Networks



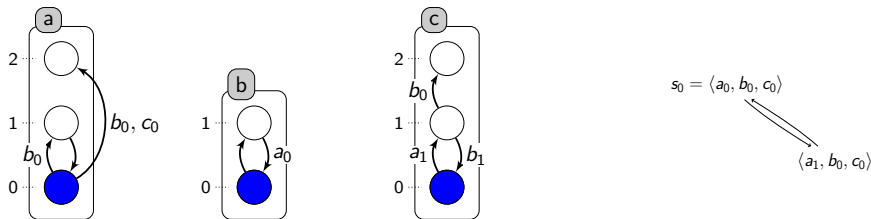
**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

**States:** sets of active local states  $\langle a_1, b_0, c_0 \rangle$

**Transitions:** dynamics  $t_1 = a_0 \xrightarrow{b_0} a_1, t_2 = a_1 \rightarrow a_0, t_3 = a_0 \xrightarrow{b_0, c_0} a_2, t_4 = b_0 \rightarrow b_1$

## Automata Networks



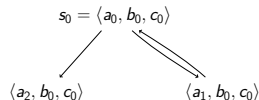
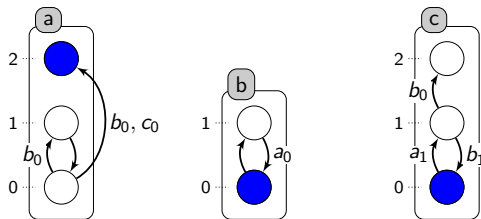
**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

**States:** sets of active local states  $\langle a_0, b_0, c_0 \rangle$

**Transitions:** dynamics  $t_1 = a_0 \xrightarrow{b_0} a_1$ ,  $t_2 = a_1 \rightarrow a_0$ ,  $t_3 = a_0 \xrightarrow{b_0, c_0} a_2$ ,  
 $t_4 = b_0 \rightarrow b_1$

## Automata Networks



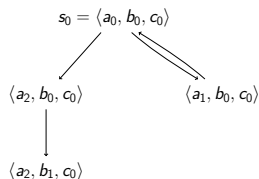
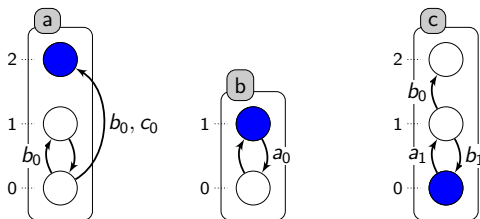
**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

**States:** sets of active local states  $\langle a_2, b_0, c_0 \rangle$

**Transitions:** dynamics  $t_1 = a_0 \xrightarrow{b_0} a_1, t_2 = a_1 \rightarrow a_0, t_3 = a_0 \xrightarrow{b_0, c_0} a_2, t_4 = b_0 \rightarrow b_1$

## Automata Networks



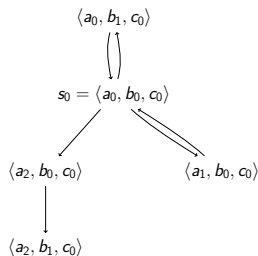
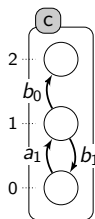
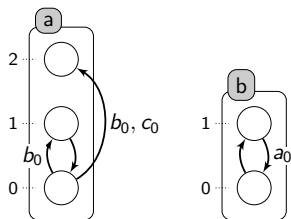
**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

**States:** sets of active local states  $\langle a_2, b_1, c_0 \rangle$

**Transitions:** dynamics  $t_1 = a_0 \xrightarrow{b_0} a_1, t_2 = a_1 \rightarrow a_0, t_3 = a_0 \xrightarrow{b_0, c_0} a_2, t_4 = b_0 \rightarrow b_1$

## Automata Networks



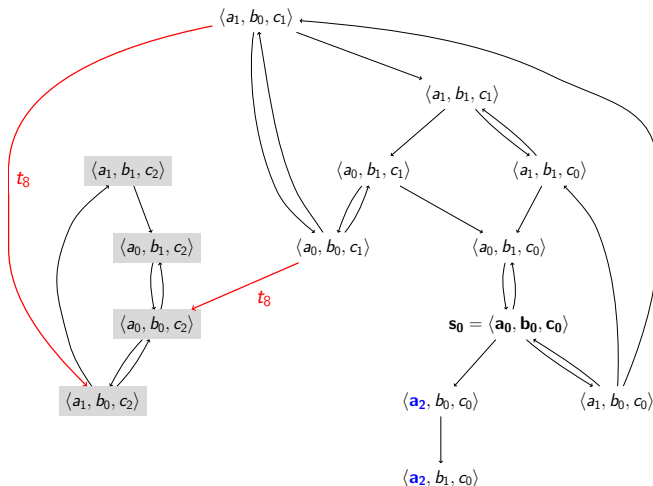
**Automata:** components  $a, b, c$

**local states:** levels of expression  $c_0, c_1, c_2$

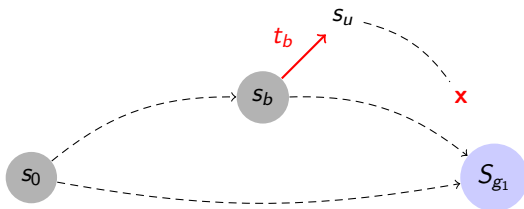
**States:** sets of active local states

**Transitions:** dynamics  $t_1 = a_0 \xrightarrow{b_0} a_1$ ,  $t_2 = a_1 \rightarrow a_0$ ,  $t_3 = a_0 \xrightarrow{b_0, c_0} a_2$ ,  
 $t_4 = b_0 \rightarrow b_1$

## Illustration of bifurcations



## Definition of bifurcation



### Definition

$t_b$  is a bifurcation transition from  $s_0$  to  $g_1$

$\iff$  there exists  $s_b$  such that:

(C1)  $s_u \not\rightarrow^* g_1$

(C2)  $s_b \rightarrow^* g_1$

(C3)  $s_0 \rightarrow^* s_b$



## Formal approximation of reachability

Static analysis by abstract interpretation

The reachability approximations for ANs introduced in [Paulevé, Magnin, Roux in *Mathematical Structures in Computer Science*, 2012]

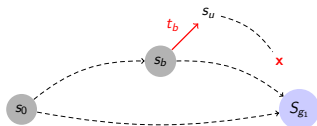
$$\text{UA}(s \rightarrow^* s') \Rightarrow s \rightarrow^* s' \Rightarrow \text{OA}(s \rightarrow^* s')$$

- OA (**over-approximations**): necessary conditions for  $s \rightarrow^* s'$
- UA (**under-approximation**): sufficient conditions for  $s \rightarrow^* s'$  (but the converse does not hold in general)

Interest:

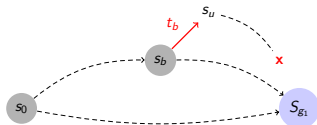
- Avoid state space explosion
- Decide in an efficient way reachability properties

## Relaxation of bifurcation problem



#: sufficient conditions.

## Relaxation of bifurcation problem



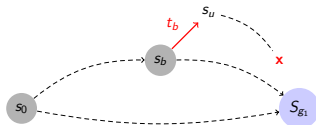
#: sufficient conditions.

$$(C1) \quad s_b \cdot t_b \not\rightarrow^* g_1$$

$$(I1^\#) \quad \neg OA(s_u \rightarrow^* g_1)$$

$$(I1^\#) \Rightarrow (C1)$$

## Relaxation of bifurcation problem



#: sufficient conditions.

$$(C1) \quad s_b \cdot t_b \not\rightarrow^* g_1$$

$$(I1^\#) \quad \neg OA(s_u \rightarrow^* g_1)$$

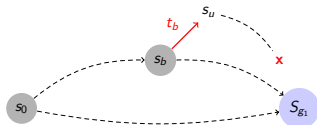
$$(I1^\#) \Rightarrow (C1)$$

$$(C2) \quad s_b \rightarrow^* g_1$$

$$(I2^\#) \quad UA(s_b \rightarrow^* g_1)$$

$$(I2^\#) \Rightarrow (C2)$$

## Relaxation of bifurcation problem



#: sufficient conditions.

$$(C1) \quad s_b \cdot t_b \not\rightarrow^* g_1$$

$$(I1^\#) \quad \neg OA(s_u \rightarrow^* g_1)$$

$$(I1^\#) \Rightarrow (C1)$$

$$(C2) \quad s_b \rightarrow^* g_1$$

$$(I2^\#) \quad UA(s_b \rightarrow^* g_1)$$

$$(I2^\#) \Rightarrow (C2)$$

$$(C3) \quad s_0 \rightarrow^* s_b$$

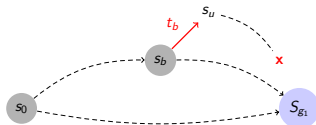
$$(I3) \quad s_b \in \text{reach}(s_0)$$

$$(I3) \Leftrightarrow (C3)$$

$$(I3^\#) \quad UA(s_0 \rightarrow^* s_b)$$

$$(I3^\#) \Rightarrow (C3)$$

## Relaxation of bifurcation problem



#: sufficient conditions.

$$(C1) \quad s_b \cdot t_b \not\rightarrow^* g_1$$

$$(I1^\#) \quad \neg OA(s_u \rightarrow^* g_1)$$

$$(I1^\#) \Rightarrow (C1)$$

$$(C2) \quad s_b \rightarrow^* g_1$$

$$(I2^\#) \quad UA(s_b \rightarrow^* g_1)$$

$$(I2^\#) \Rightarrow (C2)$$

$$(C3) \quad s_0 \rightarrow^* s_b$$

$$(I3) \quad s_b \in \text{reach}(s_0)$$

$$(I3) \Leftrightarrow (C3)$$

$$(I3^\#) \quad UA(s_0 \rightarrow^* s_b)$$

$$(I3^\#) \Rightarrow (C3)$$

$(I1^\#)$  and  $(I2^\#)$  and  $((I3) \text{ or } (I3^\#)) \Rightarrow t_b$  is a bifurcation.

## Implementation

### Formal approximation of reachability

- Analysis of local causality of transitions
- Computation of a so called **Local Causality Graph**
- OA/UA: particular patterns in LCG
- LCG size:  $\text{poly}(\#\text{automata}), \exp(|\text{single automaton}|)$

### ASP implementation

- Encode NP problem: find  $s_b, t_b$  such that

$$\neg \text{OA}(s_b \cdot t_b \rightarrow^* g_1) \text{ and } \text{UA}(s_b \rightarrow^* g_1) \text{ and } \text{UA}(s_0 \rightarrow^* s_b)$$

- enumeration with clingo

# Automata Network modelling of Biological Networks

## Transition-centered specification

- ...in opposition to function-centered of Boolean/Thomas networks
- explicit context/ causality of state changes
- closely related to (safe) Petri nets
- step semantics (purely async, purely sync, mixed)

## Modelling

- any Boolean/Thomas networks can be encoded;
- in case of logical rules uncertainty: model the union of Boolean/Thomas networks (over-approximation of behaviours)

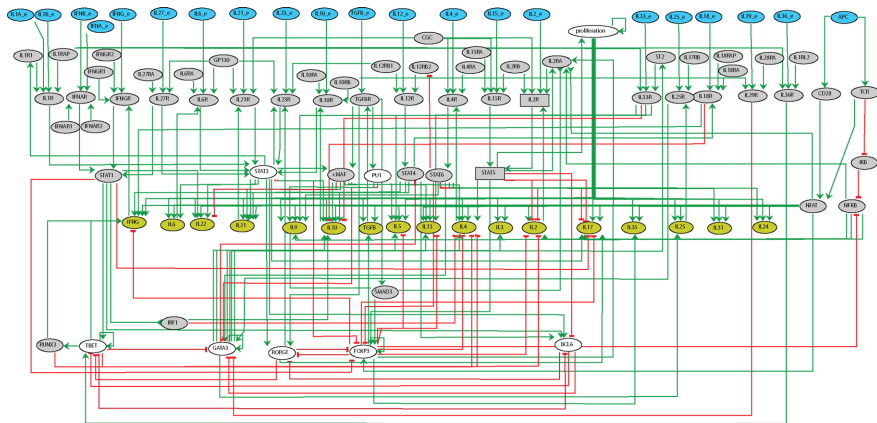


## Case study

Lambda phage : (4 components and 11 interactions);

EGF/TNF : (28 components and 55 interactions);

t\_helper differentiation: (101 components and 381 interactions).



[Abou-Jaoudé et al, in *Frontiers in Bioengineering and Biotechnology*, 2015]

## Results of identification of bifurcations

Automata Network	Goal	M-C (NuSMV)		with (I3)		with (I3 <sup>#</sup> )	
		$ t_b $	Time	$ t_b $	Time	$ t_b $	Time
Lambda phage $ \Sigma  = 4 \quad  T  = 11$	CI <sub>2</sub>	10	0.1s	6	0.1s	0	0.2s
	Cro <sub>2</sub>	3	0.1s	3	0.1s	2	0.3s
EGF/TNF $ \Sigma  = 28 \quad  T  = 55$	NFkB <sub>0</sub>	5	0.2s	4	0.1s	2	0.1s
	IKB <sub>1</sub>	5	0.2s	3	0.1s	2	0.1s
Th_th17 $ \Sigma  = 101 \quad  T  = 381$	RORGT <sub>1</sub>	18	48s	9	23s	8	26s
	BCL6 <sub>1</sub>	7	26s	5	23s	4	24s
Th_HTG $ \Sigma  = 101 \quad  T  = 381$	BCL6 <sub>1</sub>	out-of-time		out-of-time		6	61s
	GATA3 <sub>1</sub>					7	34s

Implemented in ASP (Answer Set Programming) and solve with clingo 4.5.4.

# Conclusions & Perspectives

## Summary



$$\left\{ \begin{array}{l} \text{AI (Artificial intelligence)} \\ + \\ \text{AI (Abstract interpretation)} \end{array} \right. \implies \text{formal approximation of bifurcations}$$

- Tractable on large networks (compared to model-checking)
- Under-approximation: some bifurcations are not returned.

## Perspectives

- Over-approximation of bifurcations
- Use bifurcations for the analysis of probability
- Applications for predicting targets for cellular reprogramming.

**Constraint-Based Methods for Bioinformatics Workshop**

# **Identification of Bifurcations in Biological Regulatory Networks using Answer-Set Programming**

**Louis Fippo Fitime<sup>1</sup>**

`louis.fippo-fitime@irccyn.ec-nantes.fr`

**Joint work with:**

**Olivier Roux<sup>1</sup>, Carito Guziolowski<sup>1</sup> and Loïc Paulevé<sup>2</sup>**

<sup>1</sup> MeForBio / IRCCyN / École Centrale de Nantes (Nantes, France)

<sup>2</sup> Bioinfo / LRI / CNRS (Paris, France)

`olivier.roux@irccyn.ec-nantes.fr`

`carito.guziolowski@irccyn.ec-nantes.fr`

`loic.pauleve@lri.fr`

Thank you for your attention!