Constraint-Based Methods for Bioinformatics Workshop

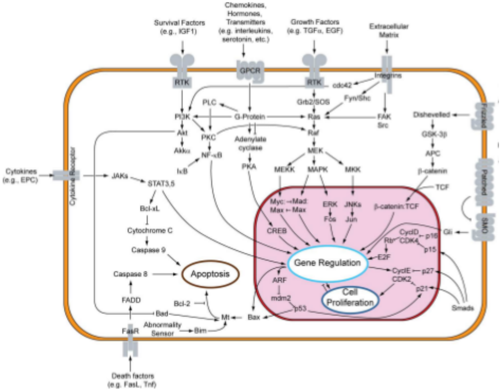
Identification of Bifurcations in Biological Regulatory Networks using Answer-Set Programming

Louis Fippo Fitime[1]
louis.fippo-fitime@irccyn.ec-nantes.fr

Joint work with:
Olivier Roux[1], Carito Guziolowski[1] and Loïc Paulevé[2]

[1] MeForBio / IRCCyN / École Centrale de Nantes (Nantes, France)
[2] Bioinfo / LRI / CNRS (Paris, France)
olivier.roux@irccyn.ec-nantes.fr
carito.guziolowski@irccyn.ec-nantes.fr
loic.pauleve@lri.fr

Context

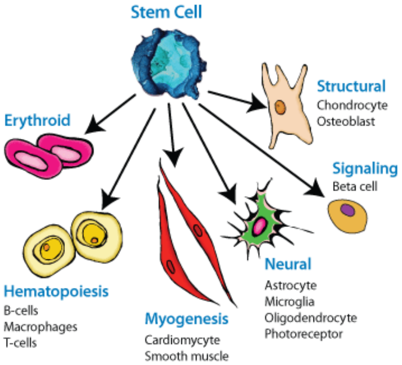


[Liu et al, in *Journal of Bioinformatics and Computational Biology*, 2012]

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

Motivation

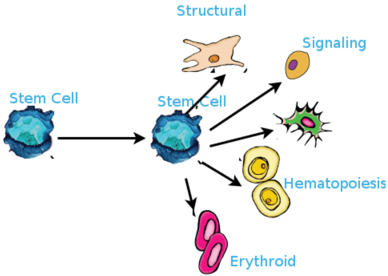
stem cell differentiation



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

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

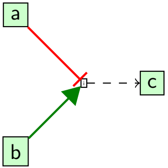
Contribution

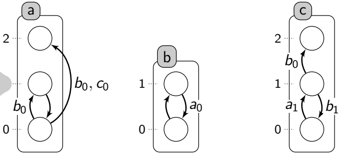


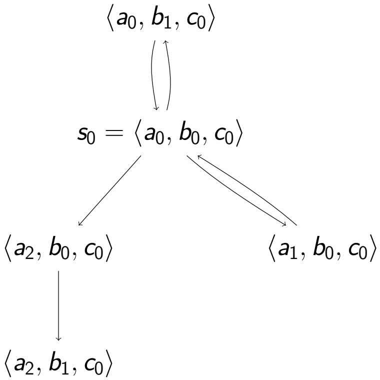
- Bifurcation transitions can be expressed in temporal logic formula
- **"Relaxation"** :
  verification of CTL formula (PSPACE (complete))
  $\Rightarrow$ NP problem which can be easily expressed in SAT/ASP
- **Contribution** :

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

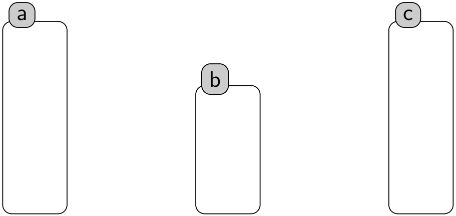
Biological networks modelling





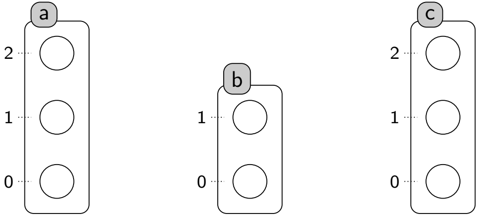


Automata Networks



Automata: components a, b, c

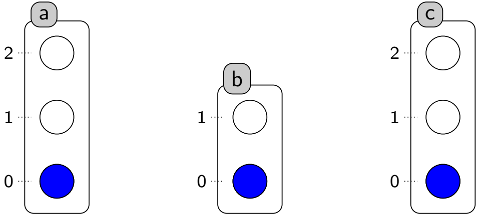
Automata Networks



Automata: components a, b, c

local states: levels of expression c_0, c_1, c_2

Automata Networks

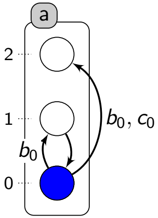


Automata: components a, b, c

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



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

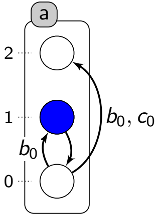
Automata: components $a,\ b,\ c$

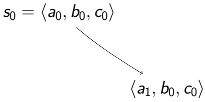
local states: levels of expression $c_0,\ c_1,\ c_2$

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





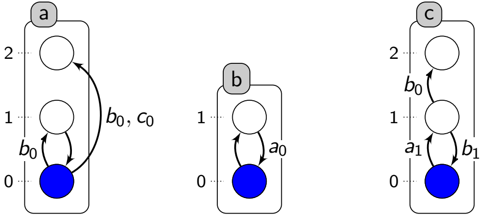
Automata: components $a,\ b,\ c$

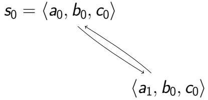
local states: levels of expression $c_0,\ c_1,\ c_2$

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

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

Automata Networks





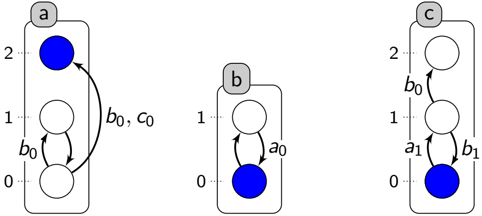
Automata: components $a,\ b,\ c$

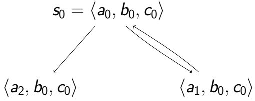
local states: levels of expression $c_0,\ c_1,\ c_2$

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

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

Automata Networks





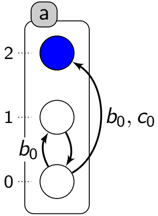
Automata: components $a,\ b,\ c$

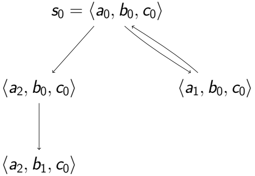
local states: levels of expression $c_0,\ c_1,\ c_2$

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

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

Automata Networks





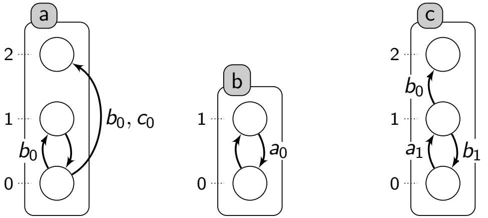
Automata: components $a,\ b,\ c$

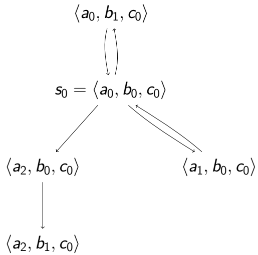
local states: levels of expression $c_0,\ c_1,\ c_2$

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





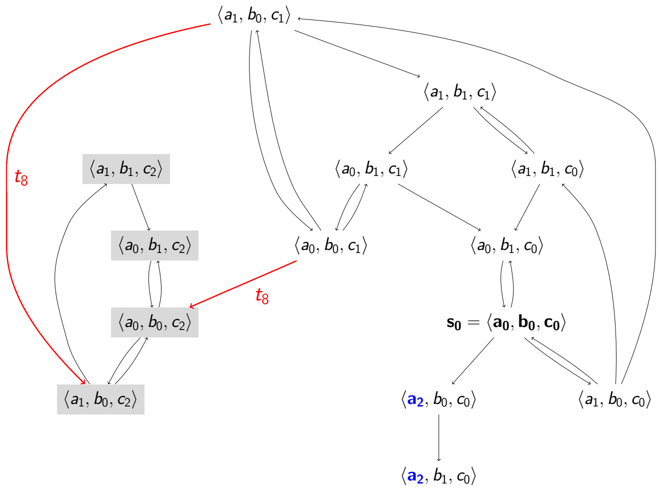
Automata: components a, b, c

local states: levels of expression c_0, c_1, c_2

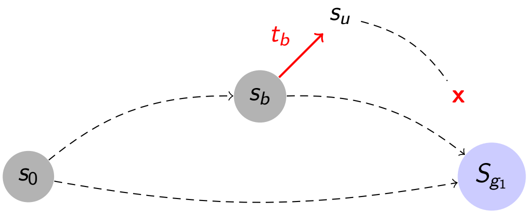
States: sets of active local states

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) $s_u \not\to^* g_1$ (C2) $s_b \to^* g_1$ (C3) $s_0 \to^* s_b$

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]

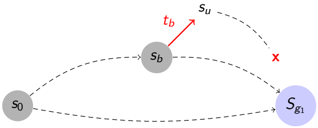
$$\mathrm{UA}(s \rightarrow^* s') \Rightarrow \mathbf{s} \rightarrow^* \mathbf{s}' \Rightarrow \mathrm{OA}(s \rightarrow^* s')$$

- 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:

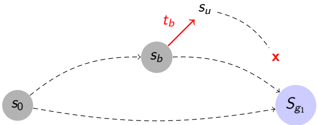
- Avoid state space explosion
- Decide in an efficient way reachability properties

Relaxation of bifurcation problem



#: sufficient conditions.

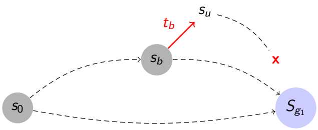
Relaxation of bifurcation problem



#: sufficient conditions.

(C1) $s_b \cdot t_b \not\to^* g_1$ (I1#) $\neg\,\mathrm{OA}(s_u \to^* g_1)$ (I1#) $\Rightarrow$ (C1)

Relaxation of bifurcation problem

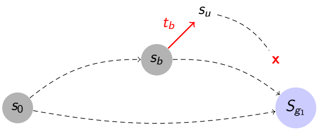


#: sufficient conditions.

(C1) $s_b \cdot t_b \not\rightarrow^* g_1$ (I1#) $\neg\,\mathrm{OA}(s_u \rightarrow^* g_1)$ (I1#) $\Rightarrow$ (C1)

(C2) $s_b \rightarrow^* g_1$ (I2#) $\mathrm{UA}(s_b \rightarrow^* g_1)$ (I2#) $\Rightarrow$ (C2)

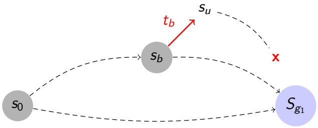
Relaxation of bifurcation problem



#: sufficient conditions.

(C1) $s_b \cdot t_b \not\to^* g_1$	(I1$^\#$) $\neg\,\mathrm{OA}(s_u \to^* g_1)$	(I1$^\#$) $\Rightarrow$ (C1)
(C2) $s_b \to^* g_1$	(I2$^\#$) $\mathrm{UA}(s_b \to^* g_1)$	(I2$^\#$) $\Rightarrow$ (C2)
(C3) $s_0 \to^* s_b$	(I3) $s_b \in \mathrm{reach}(s_0)$	(I3) $\Leftrightarrow$ (C3)
	(I3$^\#$) $\mathrm{UA}(s_0 \to^* s_b)$	(I3$^\#$) $\Rightarrow$ (C3)

Relaxation of bifurcation problem



#: sufficient conditions.

(C1) $s_b \cdot t_b \not\to^* g_1$	(I1#) $\neg\,\mathrm{OA}(s_u \to^* g_1)$	(I1#) $\Rightarrow$ (C1)
(C2) $s_b \to^* g_1$	(I2#) $\mathrm{UA}(s_b \to^* g_1)$	(I2#) $\Rightarrow$ (C2)
(C3) $s_0 \to^* s_b$	(I3) $s_b \in \mathrm{reach}(s_0)$	(I3) $\Leftrightarrow$ (C3)
	(I3#) $\mathrm{UA}(s_0 \to^* s_b)$	(I3#) $\Rightarrow$ (C3)

(I1#) *and* (I2#) *and* ((I3)*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: poly(#automata),exp(|single automaton|)

ASP implementation

- Encode NP problem: find s_b, t_b such that

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

- enumeration with clingo

Automata Network modelling of Biological Networks

Transition-centered specification

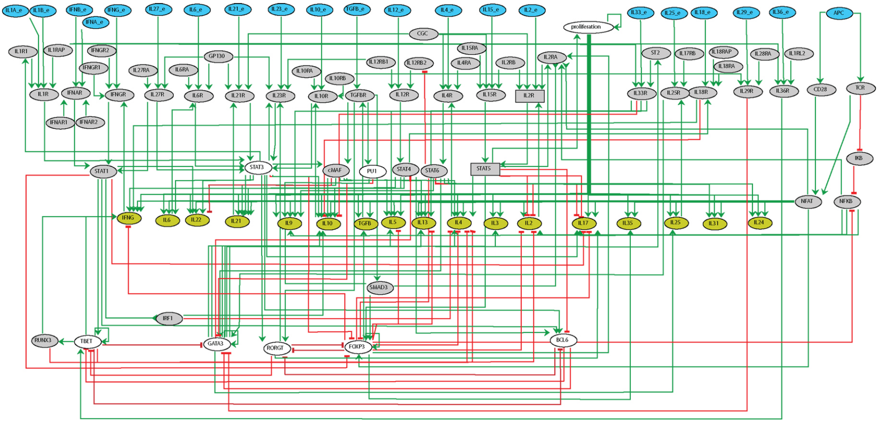
- ...in opposition to function-centered of Boolean/Thomas networks
- explicit context/ causality of state changes
- closely related to (safe) Petri nets
- step semantics (purely async, purely sync, mixed)

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 ($I3^{\#}$)	
		$\|t_b\|$	Time	$\|t_b\|$	Time	$\|t_b\|$	Time
Lambda phage	CI_2	10	$0.1s$	6	$0.1s$	0	$0.2s$
$\|\Sigma\| = 4 \quad \|T\| = 11$	Cro_2	3	$0.1s$	3	$0.1s$	2	$0.3s$
EGF/TNF	NFkB_0	5	$0.2s$	4	$0.1s$	2	$0.1s$
$\|\Sigma\| = 28 \quad \|T\| = 55$	IKB_1	5	$0.2s$	3	$0.1s$	2	$0.1s$
Th_th17	RORGT_1	18	$48s$	9	$23s$	8	$26s$
$\|\Sigma\| = 101 \quad \|T\| = 381$	$\mathrm{BCL6}_1$	7	$26s$	5	$23s$	4	$24s$
Th_HTG	$\mathrm{BCL6}_1$	out-of-time		out-of-time		6	$61s$
$\|\Sigma\| = 101 \quad \|T\| = 381$	$\mathrm{GATA3}_1$					7	$34s$

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

Conclusions & Perspectives

Summary

- $$\left\{\begin{array}{l}\text{AI (Artificial inteligence)}\\ +\\ \text{AI (Abstract interpretation)}\end{array}\right. \Longrightarrow \text{formal approximation of bifurcations}$$
- 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[1]
louis.fippo-fitime@irccyn.ec-nantes.fr

Joint work with:
Olivier Roux[1], Carito Guziolowski[1] and Loïc Paulevé[2]

[1] MeForBio / IRCCyN / École Centrale de Nantes (Nantes, France)
[2] 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!