

Soutenance de thèse de doctorat

Modélisation hybride, Analyse et Vérification des réseaux de régulation biologique

Louis FIPPO FITIME

louis.fippo-fitime@ircbyn.ec-nantes.fr

<http://www.ircbyn.ec-nantes.fr/~fippofit/>

Lundi 28 novembre 2016

Rapporteurs: Christine FROIDEVAUX, Professeur des universités, Université Paris Sud
Simon de GIVRY, Chargé de recherche INRA, INRA Toulouse

Examinateurs: Anne POUPON, Directrice de recherche CNRS, INRA de Tours
Sylvain SOLIMAN, Chargé de recherche INRIA, INRIA Paris Saclay

Directeur de thèse: Olivier ROUX, Professeur des universités, École centrale de Nantes

Co-encadrant de thèse: Carito GUZIOLOWSKI, Maître de conférences, École centrale de Nantes

- Introduction

Contexte

Systèmes **biologiques**

Un système biologique est un ensemble d'organes/composants interagissant au sein d'un organisme dans la réalisation d'une fonction biologique commune.

- Introduction

Contexte

Systèmes complexes

Systèmes **biologiques**

Systèmes complexes sont des réseaux des composants inter-agissants entre eux,
généralement de façon non déterministe.

- Introduction

Contexte

Systèmes complexes

Systèmes **biologiques**

+

Méthodes **informatiques**

Méthodes informatiques: traitement automatique de l'information.

- Introduction

Contexte

Systèmes complexes

Systèmes **biologiques**

+

Méthodes **informatiques**

Modéliser

Un modèle est une représentation simplifiée d'un système.

- Introduction

Contexte

Systèmes complexes

Systèmes **biologiques**

+

Méthodes **informatiques**

Analyser

Modéliser

Une analyse est une étude minutieuse, précise faite pour comprendre un système (modèle), pour l'expliquer, l'éclairer.

- Introduction

Contexte

Systèmes complexes

Systèmes **biologiques**

+

Méthodes **informatiques**

Analyser

Modéliser

Vérifier

Une vérification (formelle) est le fait de prouver (de façon formelle) la présence ou non d'une propriété dans un système.

- Introduction

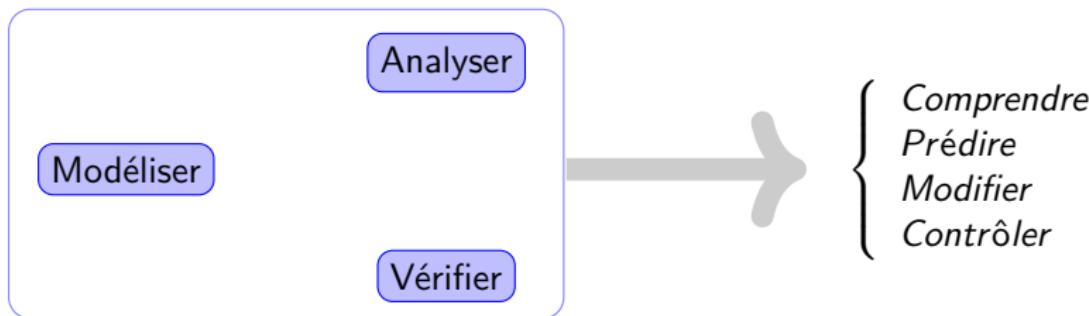
Contexte

Systèmes complexes

Systèmes **biologiques**

+

Méthodes **informatiques**



- Introduction

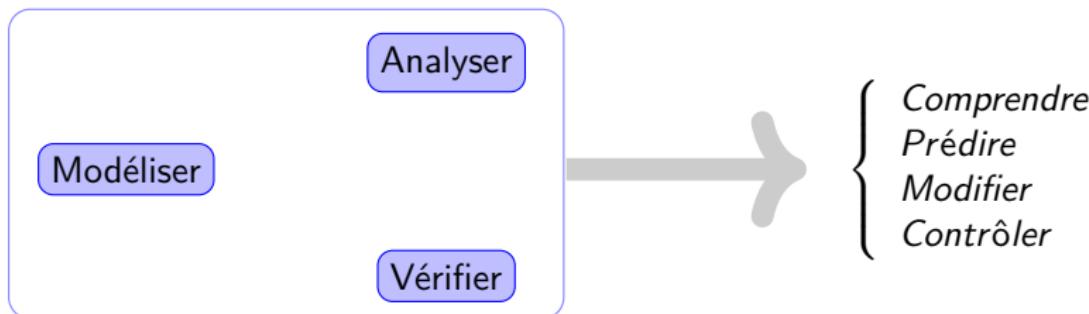
Contexte

Systèmes complexes

Systèmes **biologiques**

+

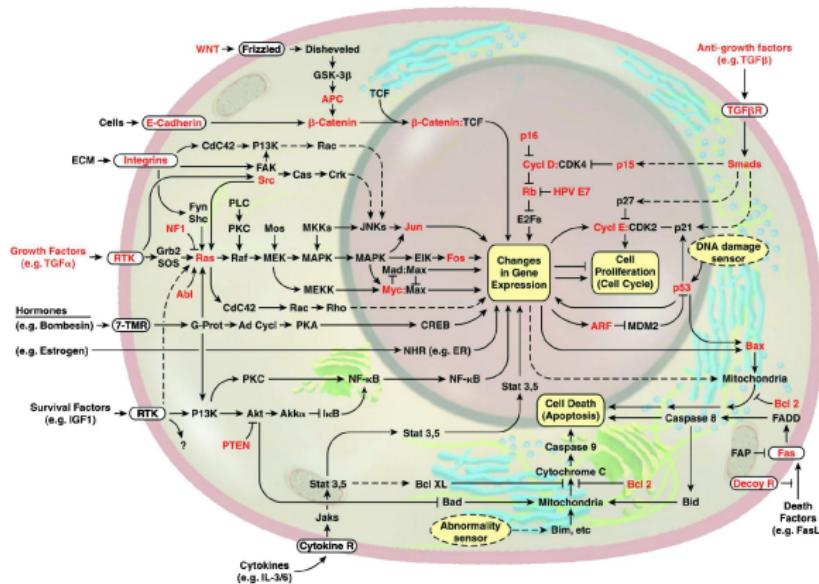
Méthodes **informatiques**



But: Développer les méthodes efficaces (complexité de calcul)
pour l'analyse des grands modèles.

- Introduction

Contexte

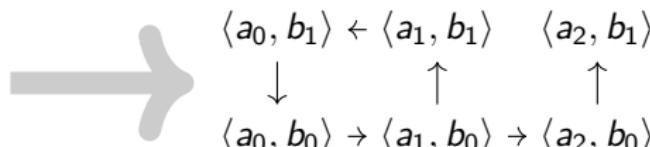
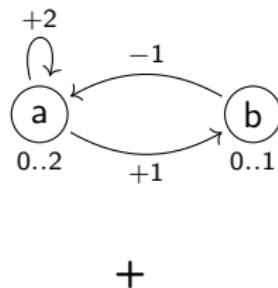


[Liu et al., in *Journal of Bioinformatics and Computational Biology*, 2012]

- Processus cellulaire \Rightarrow des réseaux de régulation biologique (RRB).
- Nœuds (les composants biologiques), les arcs (les interactions).

Modélisation discrètes

- Chaque composant possède un nombre fini de **niveaux qualitatifs** ($\{0,1,2,\dots\}$).
- Fonctions donnant le **niveau suivant** l'état des régulateurs.



Fonction de transition

[Kauffman, *Journal of Theoretical Biology*, 1969]
[Thomas, *Journal of Theoretical Biology*, 1973]

Modélisation hybride

États discrets et transitions caractéristiques hybrides

Introduction de l'aléatoire et du temps pour les transitions

Modèles stochastiques

Variables aléatoires généralement exponentielles.

- π -calcul stochastique.
- Réseaux de Petri stochastiques.

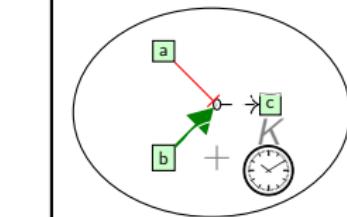
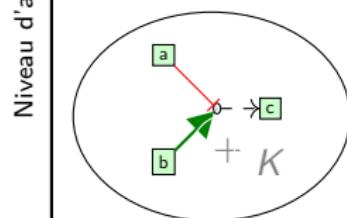
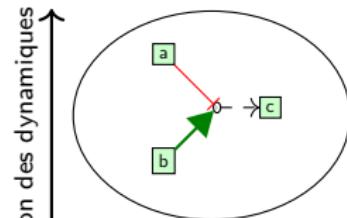
Modèles temporisés

- Automates temporisés / hybrides.
- Réseaux de Petri temporisés.

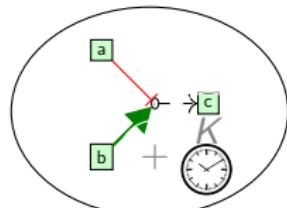
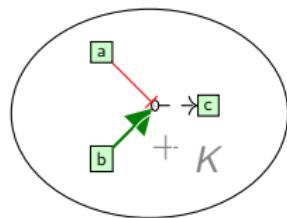
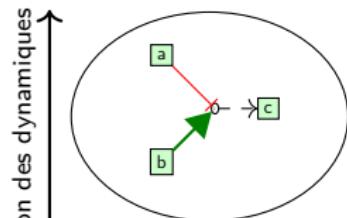
Frappes de processus.

- Contributions

Contributions



- Contributions



Contributions

Propriétés générales: bornes
sur le nbr d'attracteurs;
fonctionnalités,...

États stables

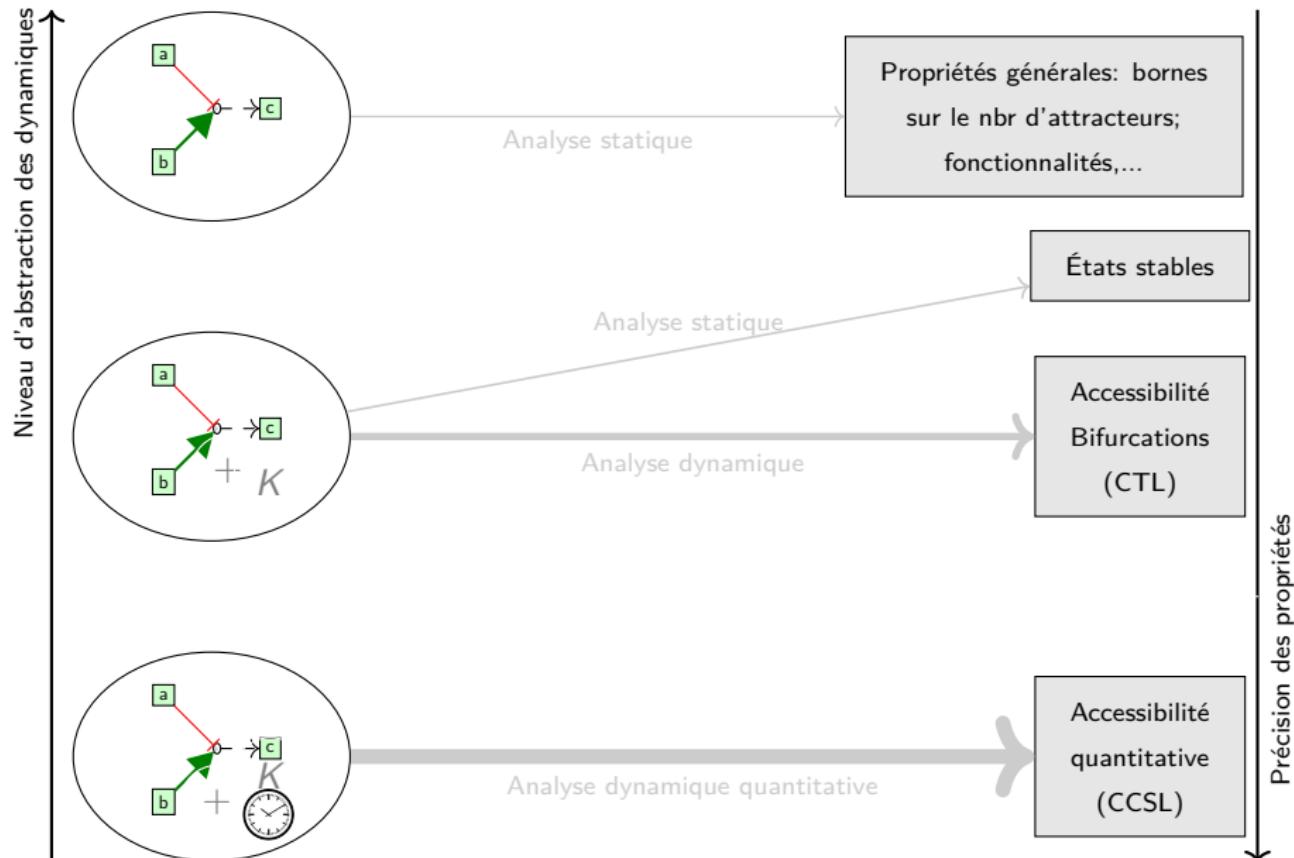
Accessibilité
Bifurcations
(CTL)

Accessibilité
quantitative
(CCSL)

Précision des propriétés

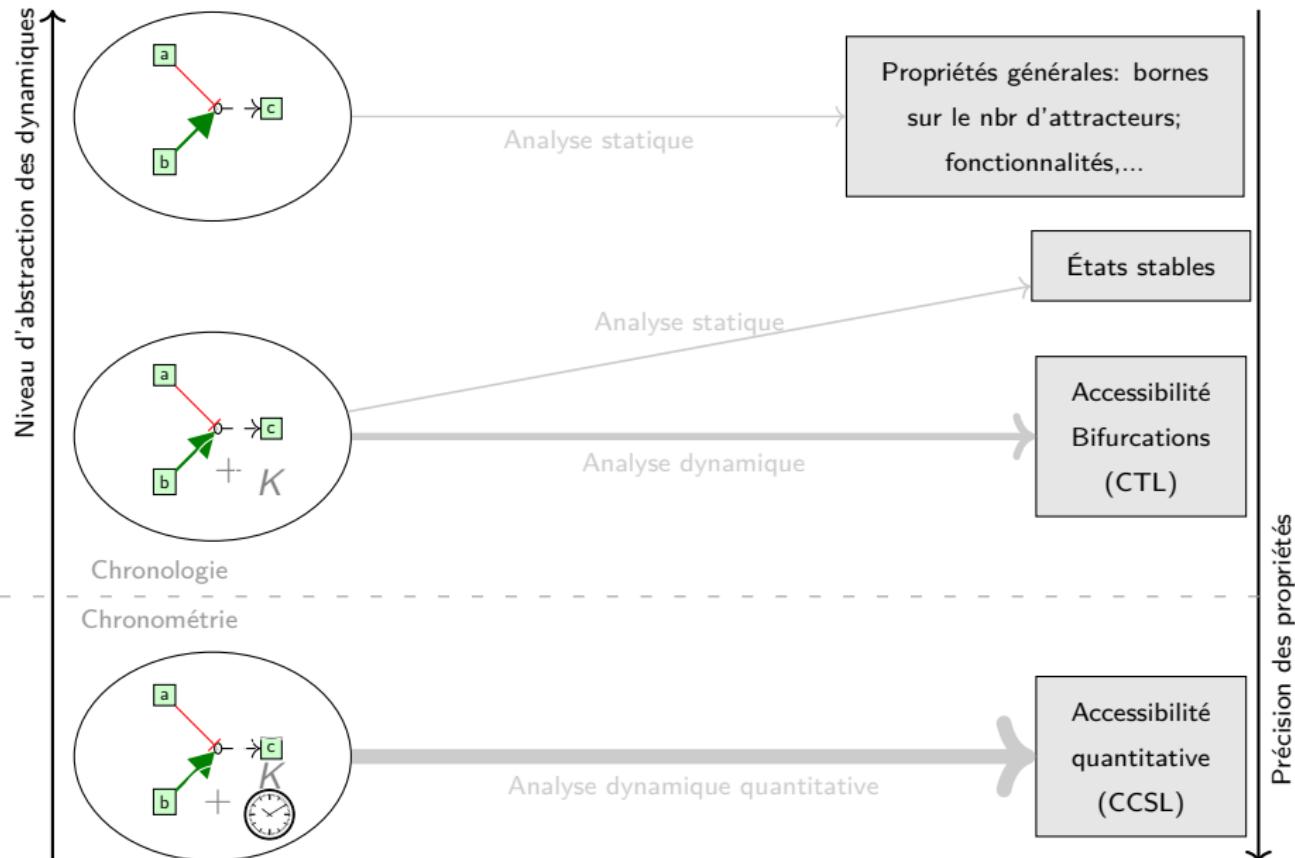
- Contributions

Contributions



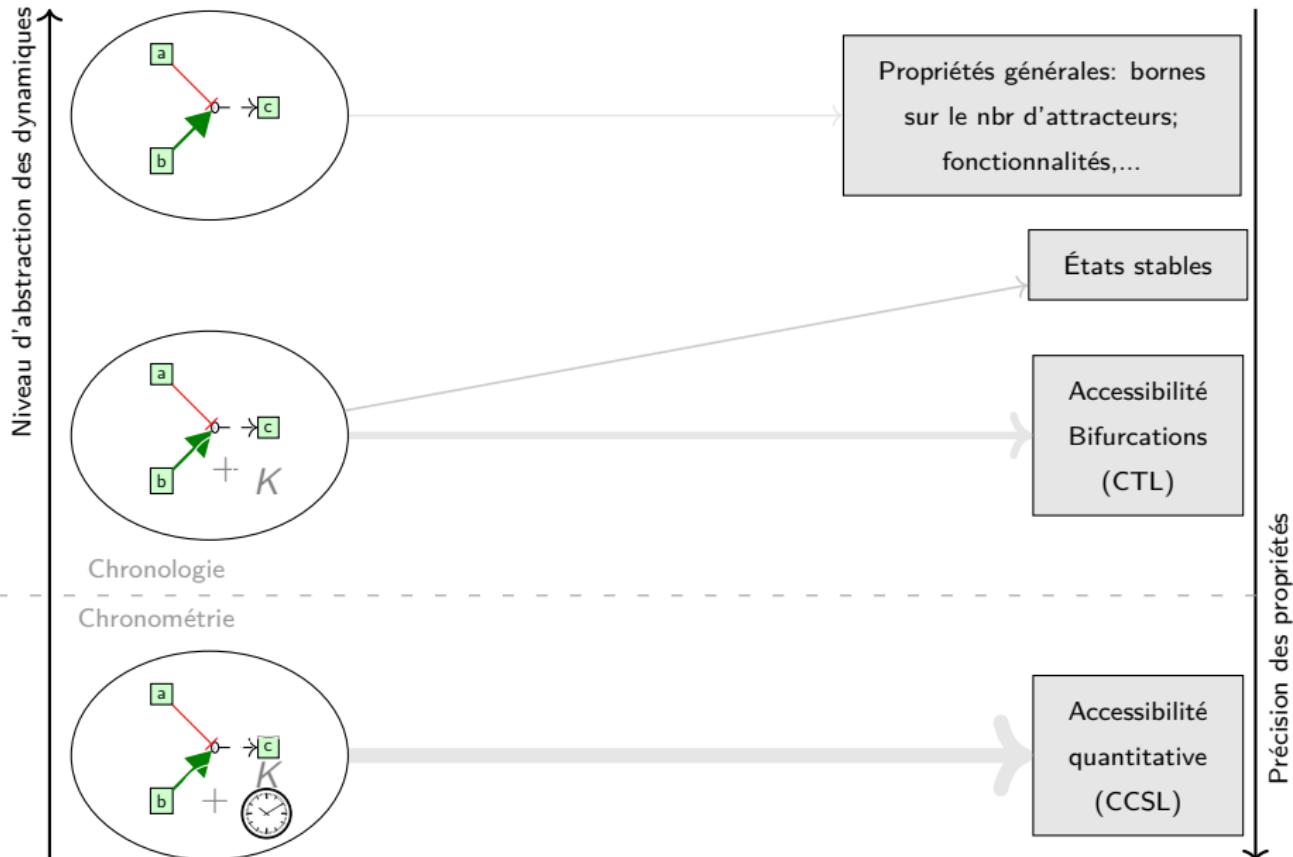
- Contributions

Contributions



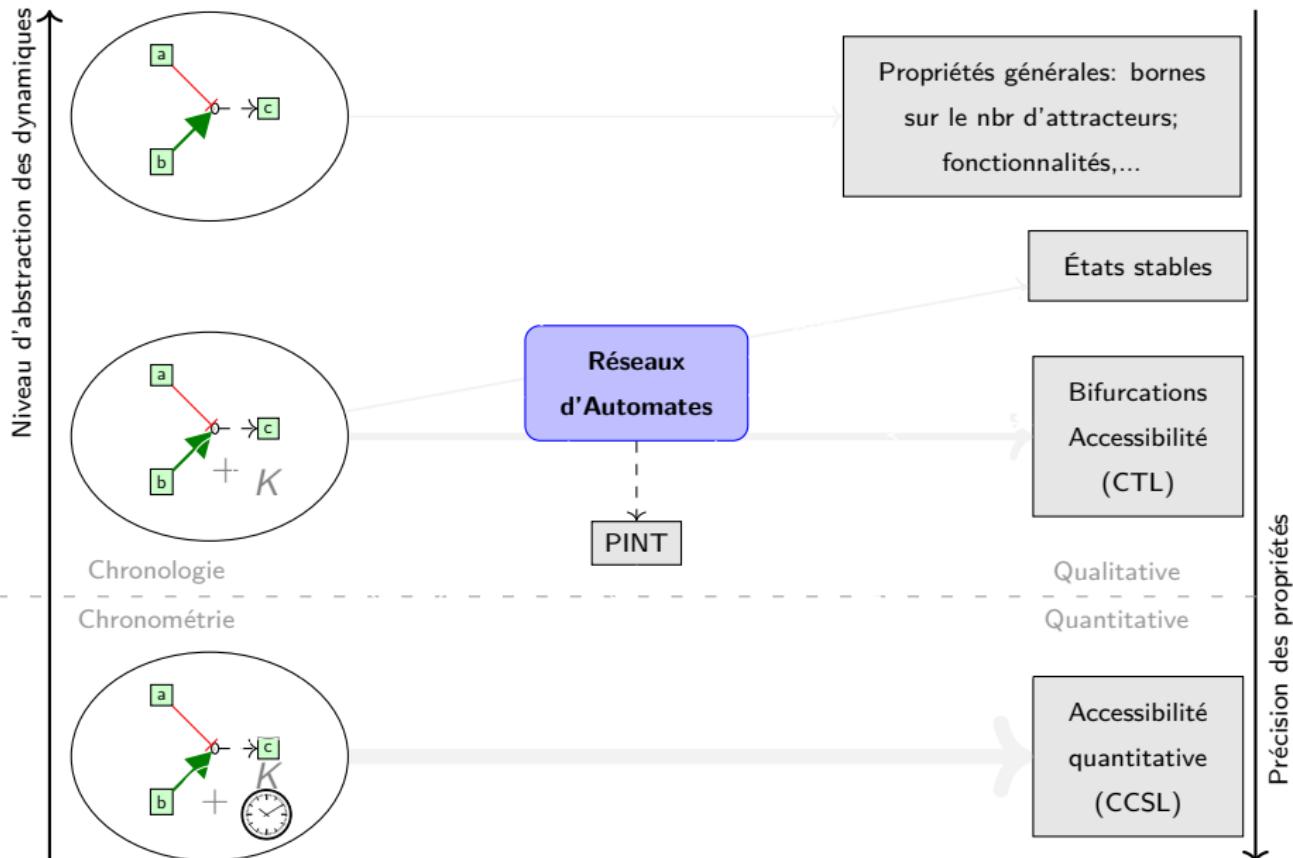
- Contributions

Contributions



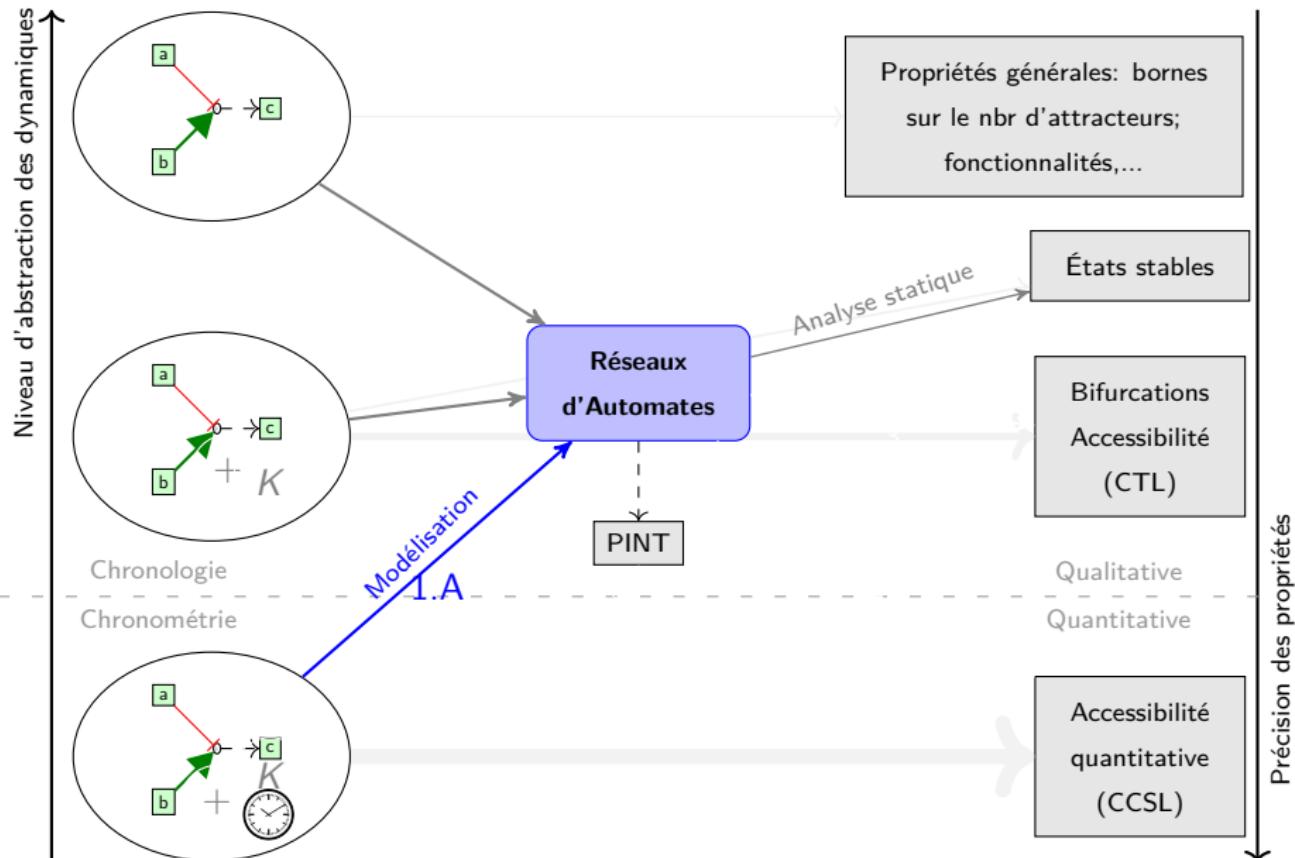
- Contributions

Contributions



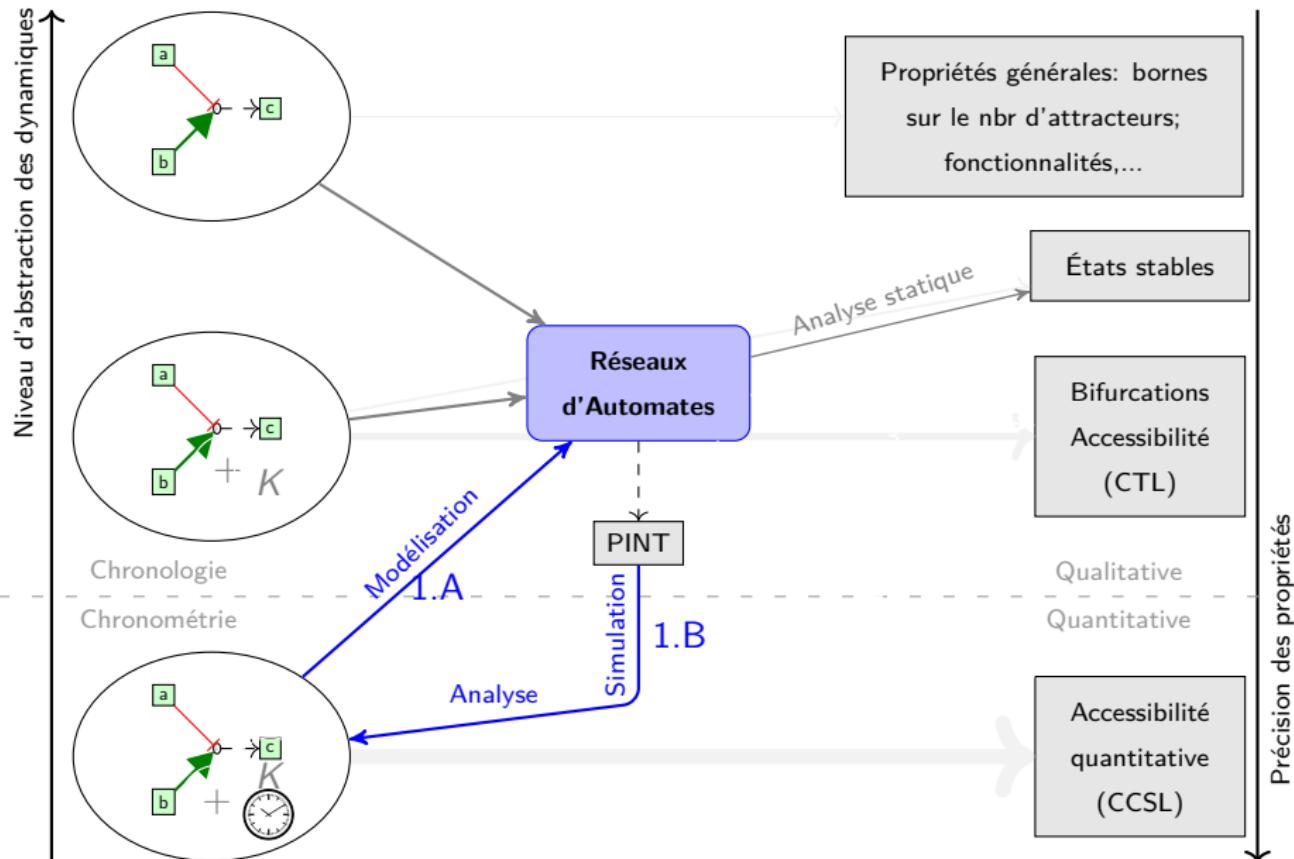
- Contributions

Contributions



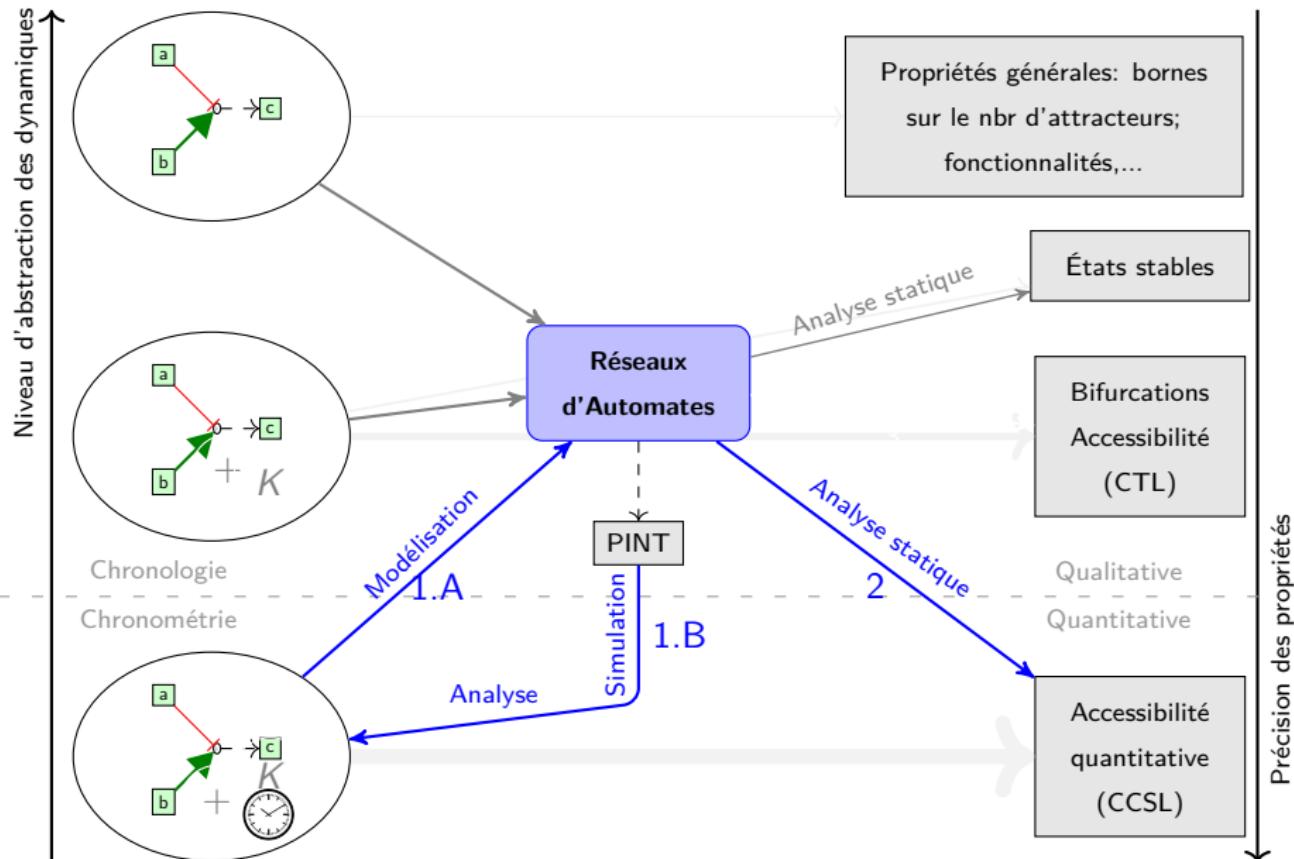
- Contributions

Contributions



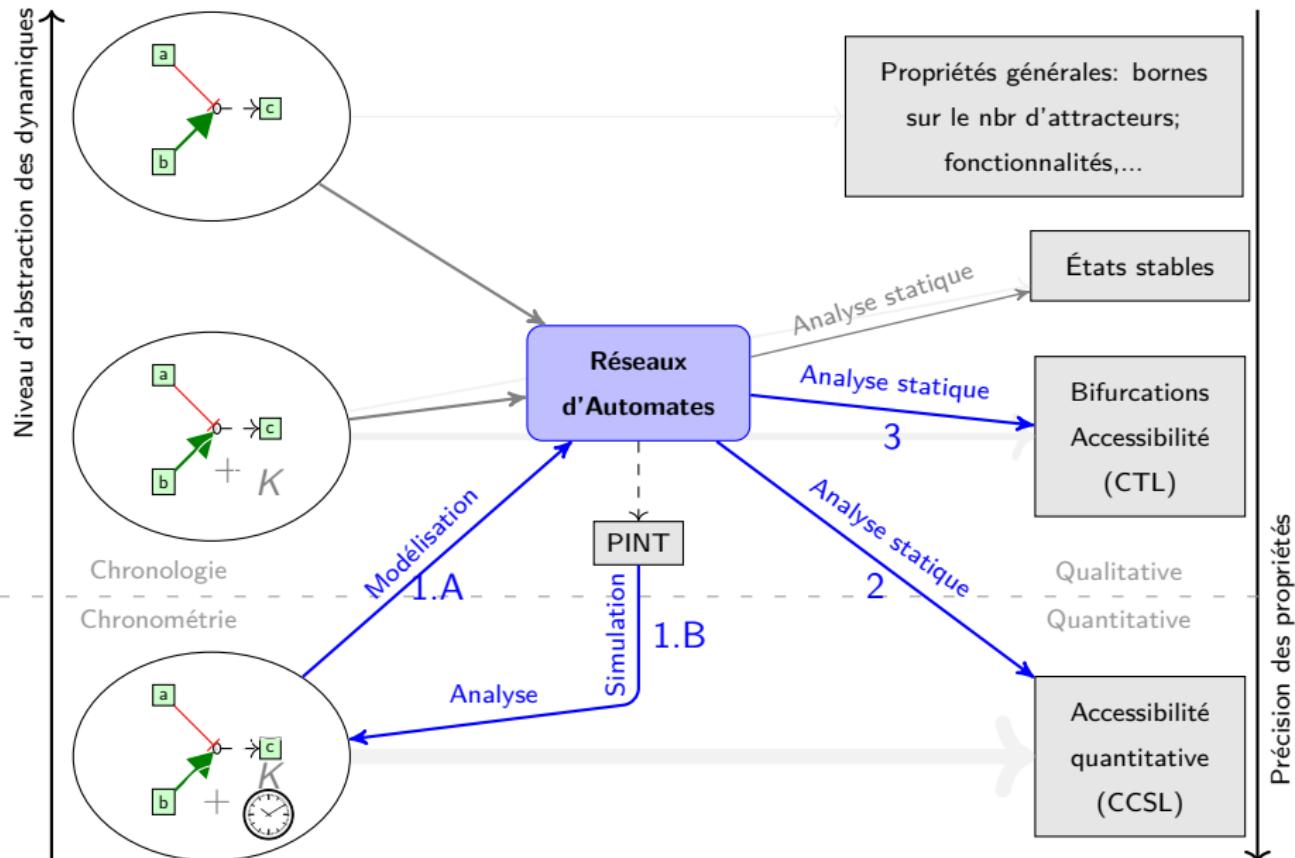
- Contributions

Contributions



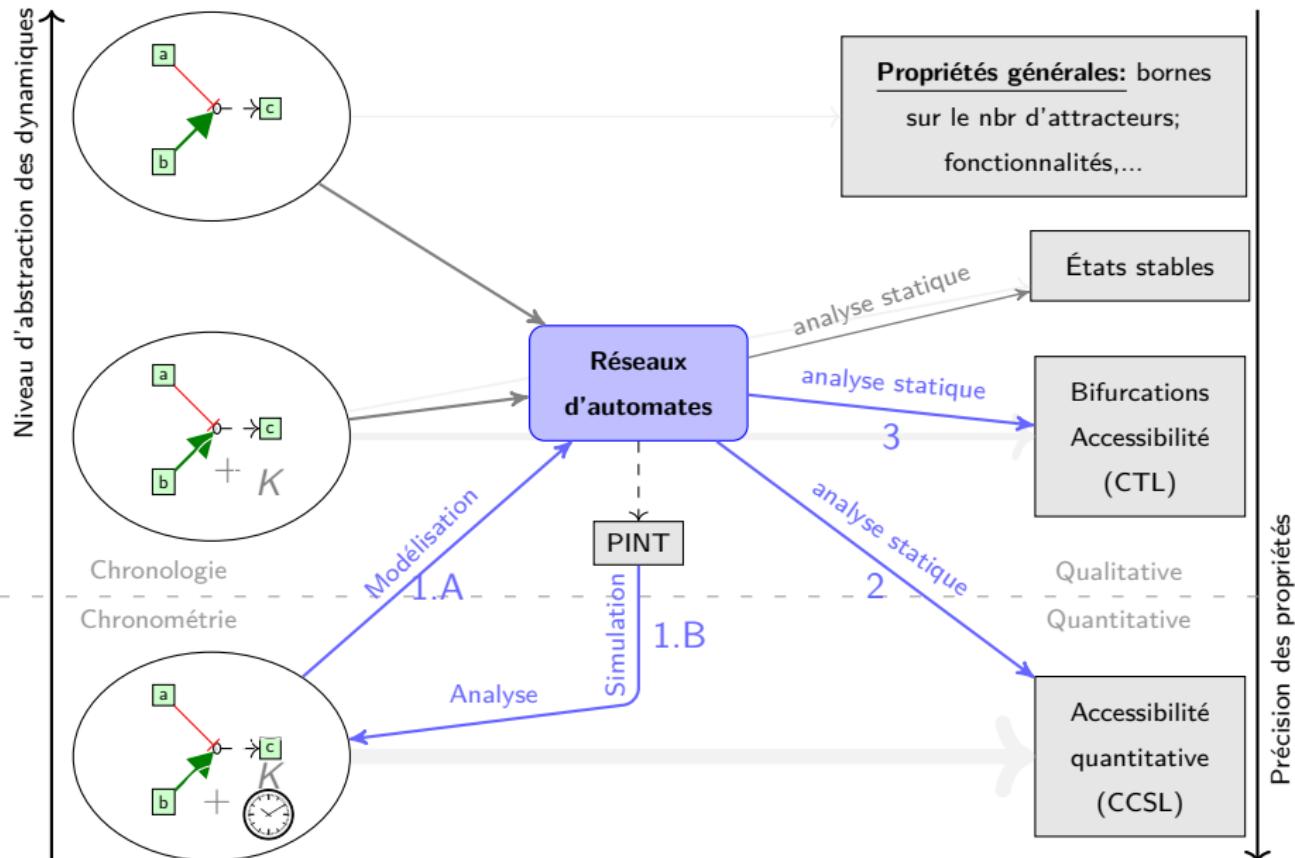
- Contributions

Contributions

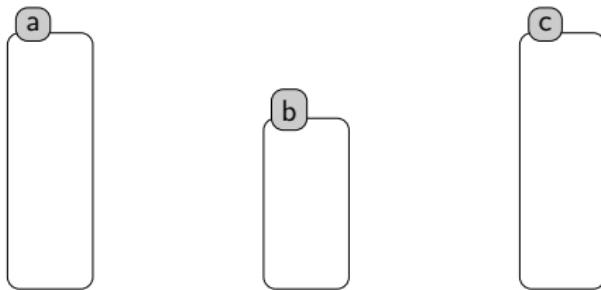


- Contributions

Contributions

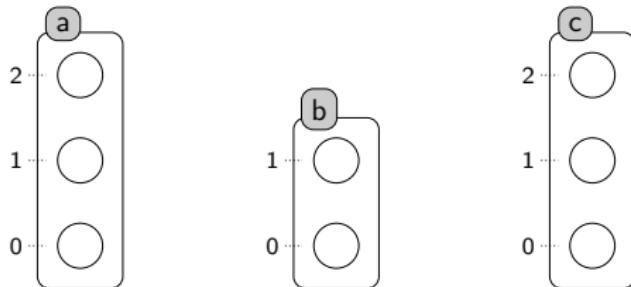


Réseaux d'automates (AN)



Automate: composants a, b, c .

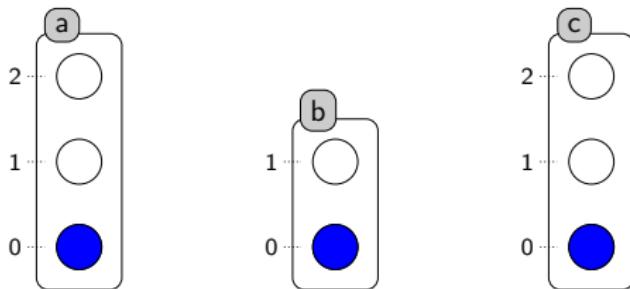
Réseaux d'automates (AN)



Automate: composants a, b, c .

États locaux: niveaux d'expression c_0, c_1, c_2 .

Réseaux d'automates (AN)

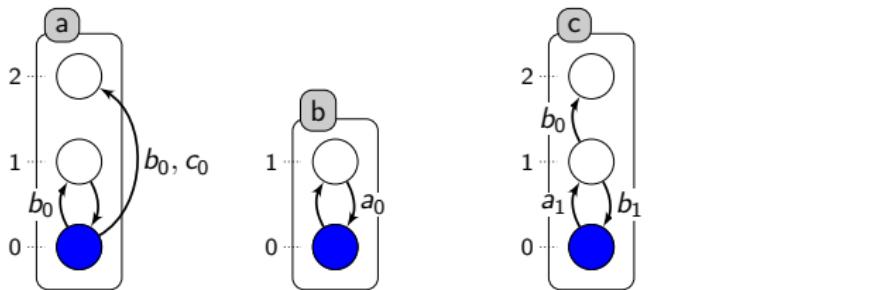


Automate: composants *a*, *b*, *c*.

États locaux: niveaux d'expression c_0, c_1, c_2 .

États: l'ensemble des états locaux actifs $\langle a_0, b_0, c_0 \rangle$

Réseaux d'automates (AN)



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

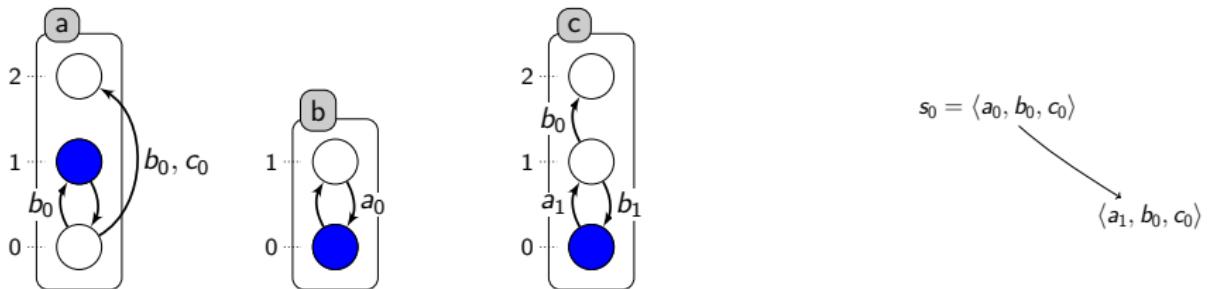
Automate: composants a, b, c .

États locaux: niveaux d'expression c_0, c_1, c_2 .

États: l'ensemble des états locaux actifs $\langle a_0, b_0, c_0 \rangle$

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

Réseaux d'automates (AN)



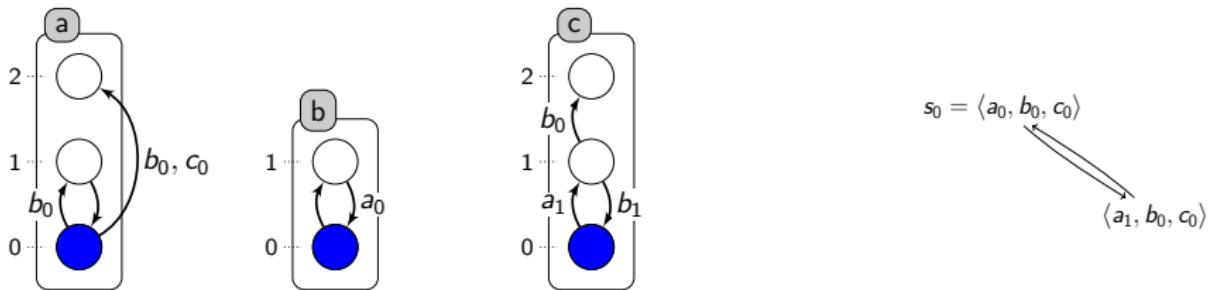
Automate: composants **a, b, c**.

États locaux: niveaux d'expression **c_0, c_1, c_2** .

États: l'ensemble des états locaux actifs $\langle a_1, b_0, c_0 \rangle$

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

Réseaux d'automates (AN)



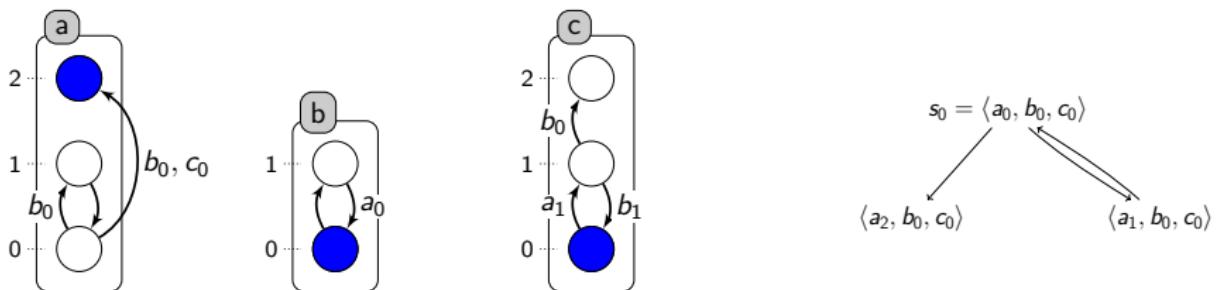
Automate: composants **a**, **b**, **c**.

États locaux: niveaux d'expression **c_0, c_1, c_2** .

États: l'ensemble des états locaux actifs $\langle a_0, b_0, c_0 \rangle$

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

Réseaux d'automates (AN)



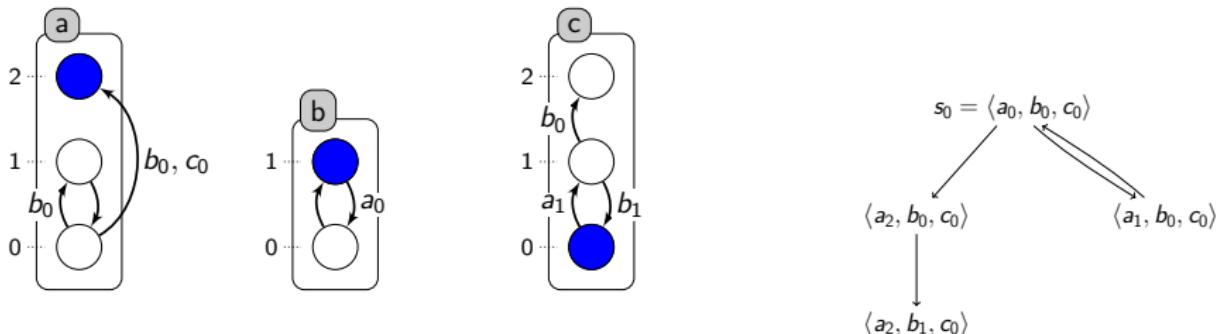
Automate: composants *a*, *b*, *c*.

États locaux: niveaux d'expression c_0, c_1, c_2 .

États: l'ensemble des états locaux actifs $\langle a_2, b_0, c_0 \rangle$

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

Réseaux d'automates (AN)



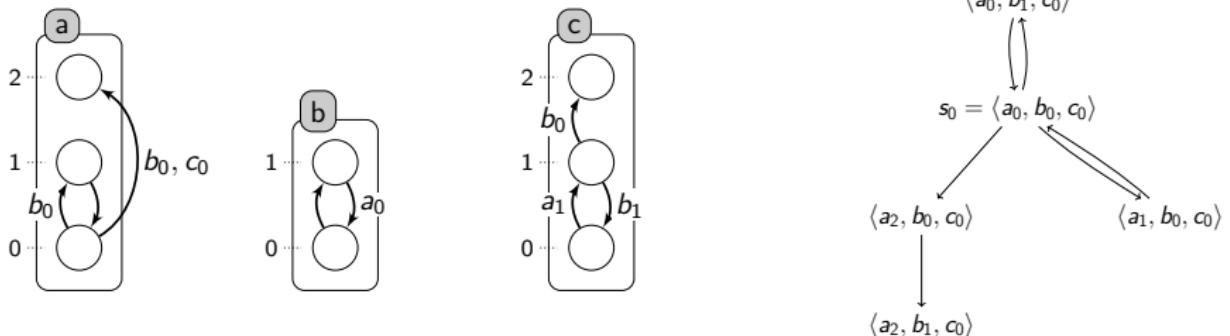
Automate: composants a, b, c .

États locaux: niveaux d'expression c_0, c_1, c_2 .

États: l'ensemble des états locaux actifs $\langle a_2, b_1, c_0 \rangle$

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

Réseaux d'automates (AN)



Automate: composants *a*, *b*, *c*.

États locaux: niveaux d'expression c_0, c_1, c_2 .

États: l'ensemble des états locaux actifs

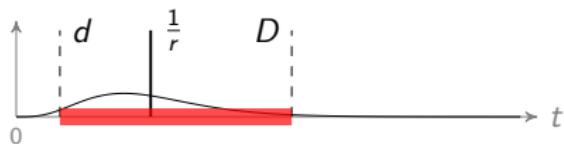
Transitions: dynamique $t_1 = a_0 \xrightarrow{b_0} a_1$, $t_2 = a_1 \rightarrow a_0$, $t_3 = a_0 \xrightarrow{b_0, c_0} a_2$, $t_4 = b_0 \rightarrow b_1$

Paramètres temporels et stochastiques

Idée: Calibrage temporel réglable avec un paramètre stochastique.

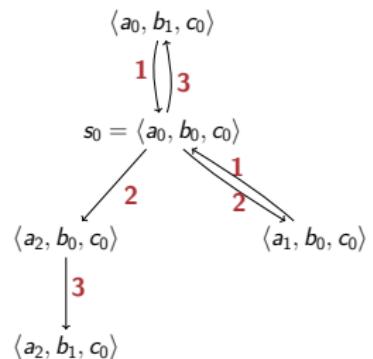
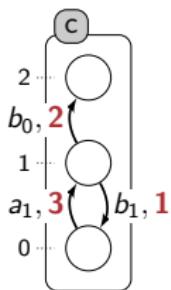
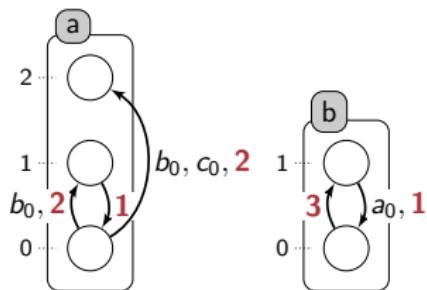
[Paulevé, Magnin, Roux in *Transactions on Computational Systems Biology*, 2011]

- Taux r & facteur d'absorption de stochasticité sa .
- Durée moyenne: r^{-1} ; variance: $r^{-2}sa^{-1}$ (distribution d'Erlang).
- Paramètres: soit (r, sa) , ou l'intervalle de tir $[d; D]$.



temps d'une action

Réseaux d'automates stochastiques



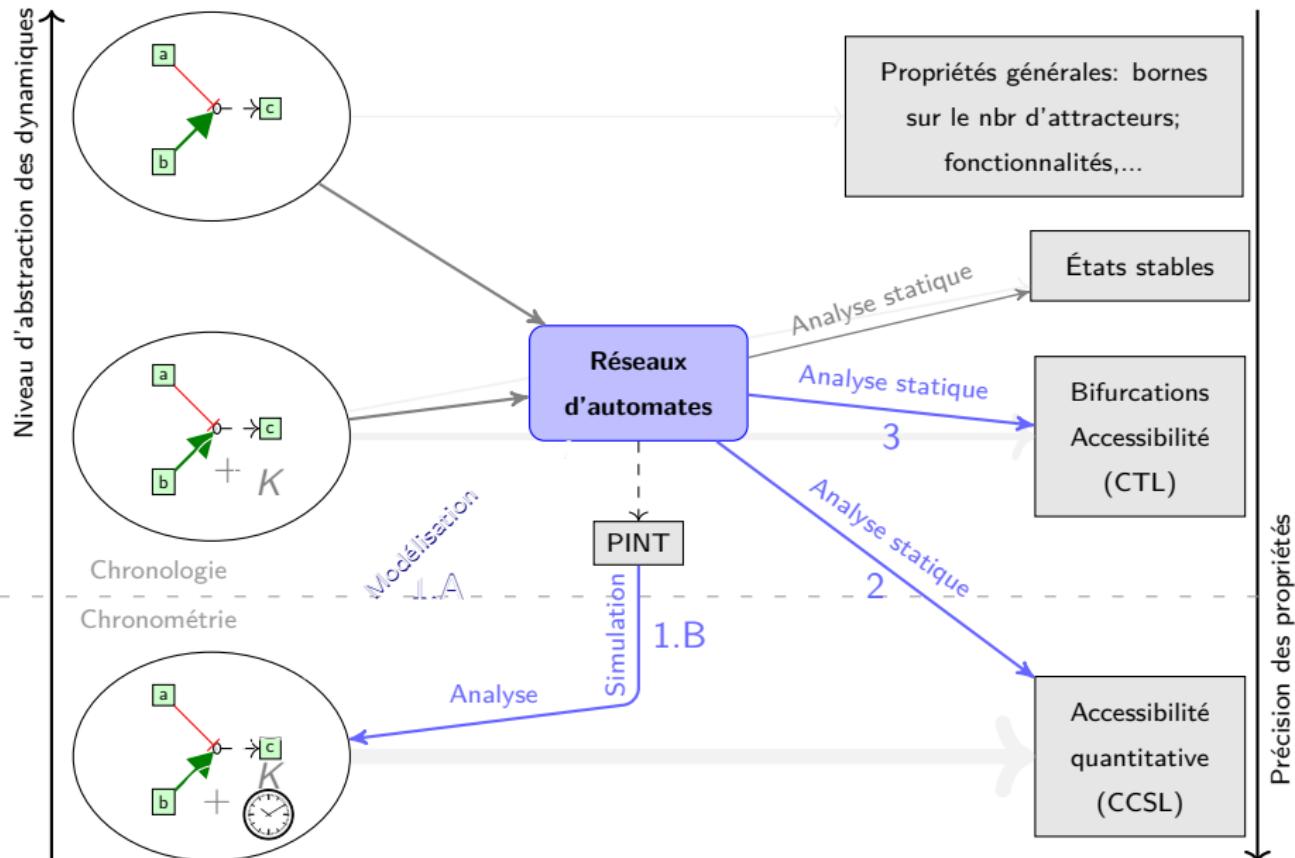
Automates: composants *a*, *b*, *c*

États locaux: niveaux d'expression *c*₀, *c*₁, *c*₂

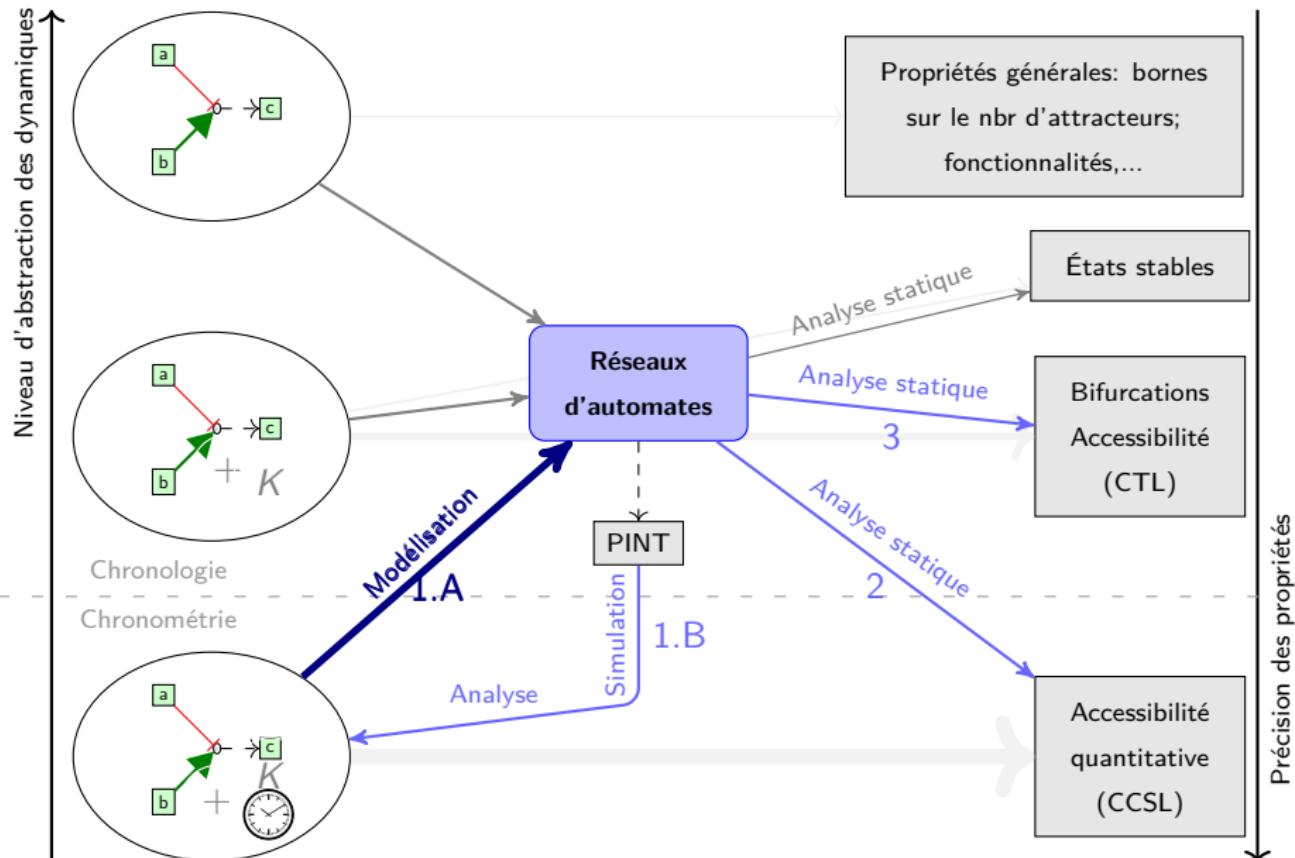
État: ensemble des états locaux actifs $\langle a_2, b_1, c_0 \rangle$

Transitions: dynamique $t_1 = a_0 \xrightarrow{b_0, 2} a_1$, $t_2 = a_1 \xrightarrow{1} a_0$, $t_3 = a_0 \xrightarrow{b_0, c_0, 2} a_2$, $t_4 = b_0 \xrightarrow{3} b_1$

Contributions



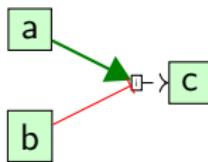
Contributions



- Raffinement de la dynamique par intégration des séries temporelles

Connaissances générales
Littérature
Hypothèses

Modélisation hybride



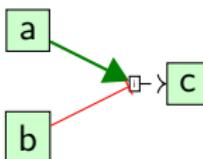
Données de séries temporelles

Genes	1h	...	24h
Gene 1		...	
Gene 2		...	

- Raffinement de la dynamique par intégration des séries temporelles

Connaissances générales
Littérature
Hypothèses

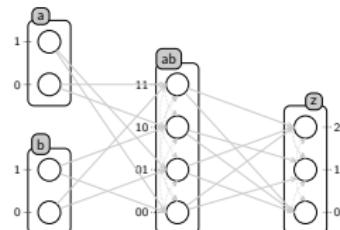
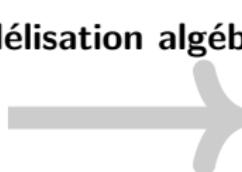
Modélisation hybride



Modélisation algébrique

Données de séries temporelles

Genes	1h	...	24h
Gene 1		...	
Gene 2		...	



- Caractérisation formelle de la topologie et de la dynamique des RRB.
- Intégration des données de séries temporelles.
- Raffinement qualitatif et quantitatif de la dynamique.
- Simulation stochastique et analyse statistique des traces.

- Raffinement de la dynamique par intégration des séries temporelles

Réseaux biologiques

Réseaux RSTC

Données expérimentales

Séries temporelles (TSD)

Modèles hybrides

Modèles AN

Estimation des paramètres

r et sa

Discrétisation

TSD

Simulation

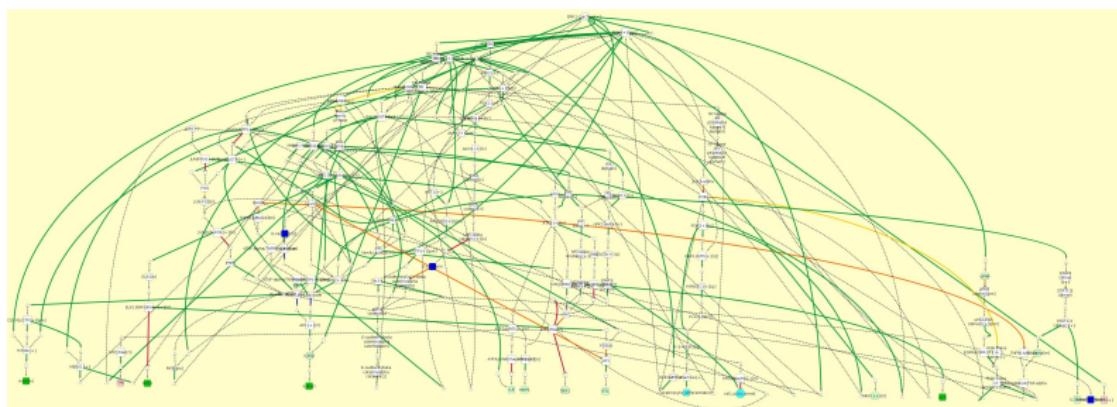
PINT

Validation du modèle

Réseau type RSTC

multi-layer Receptor-Signaling-Transcription-Cell state

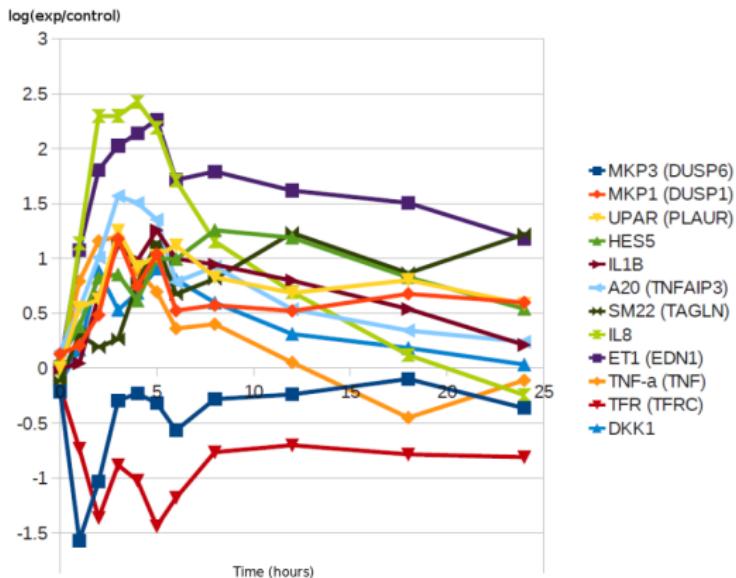
[Guziolowski et al, in *Febs Journal*, 2012]



- Cas d'étude **Ecad** extrait de Pathway Interaction Database (PID).
- **293 noeuds**: protéines de signalisation, facteurs de transcription, gènes,...
- **375 interactions**: activations, inhibitions, dissociation des complexes,...

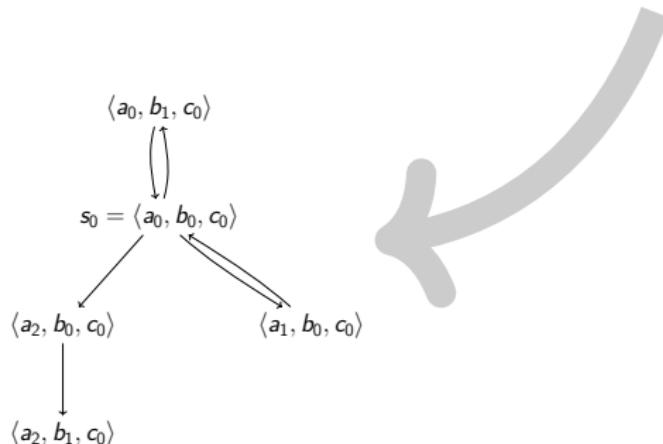
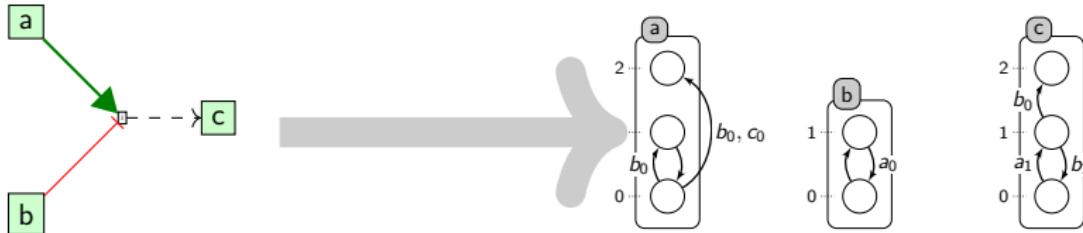
Données de séries temporelles (TSD)

[Peter Angel, Christian Schuster, DKFZ]



- Expérience: stimulation au calcium d'Ecad.
- Mesurer à 10 intervalles de temps (0-24hrs).
- Sélection de 200 transcripts (dynamique).

Modélisation des RSTC par les réseaux d'automates (AN)



Formalisation des réseaux RSTC

Détection et transformation automatique des motifs minimaux

Réseau biologique
réseau RSTC

Formalisation des réseaux RSTC

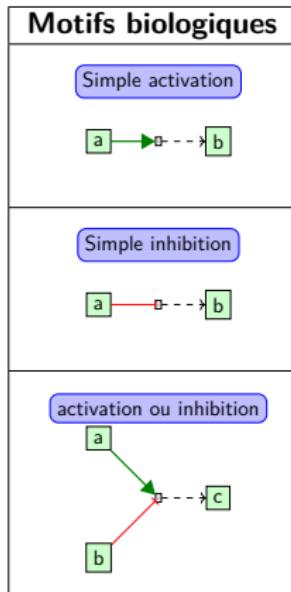
Détection et transformation automatique des motifs minimaux

Réseau biologique
réseau RSTC

Détection des motifs
dans les RSTC network

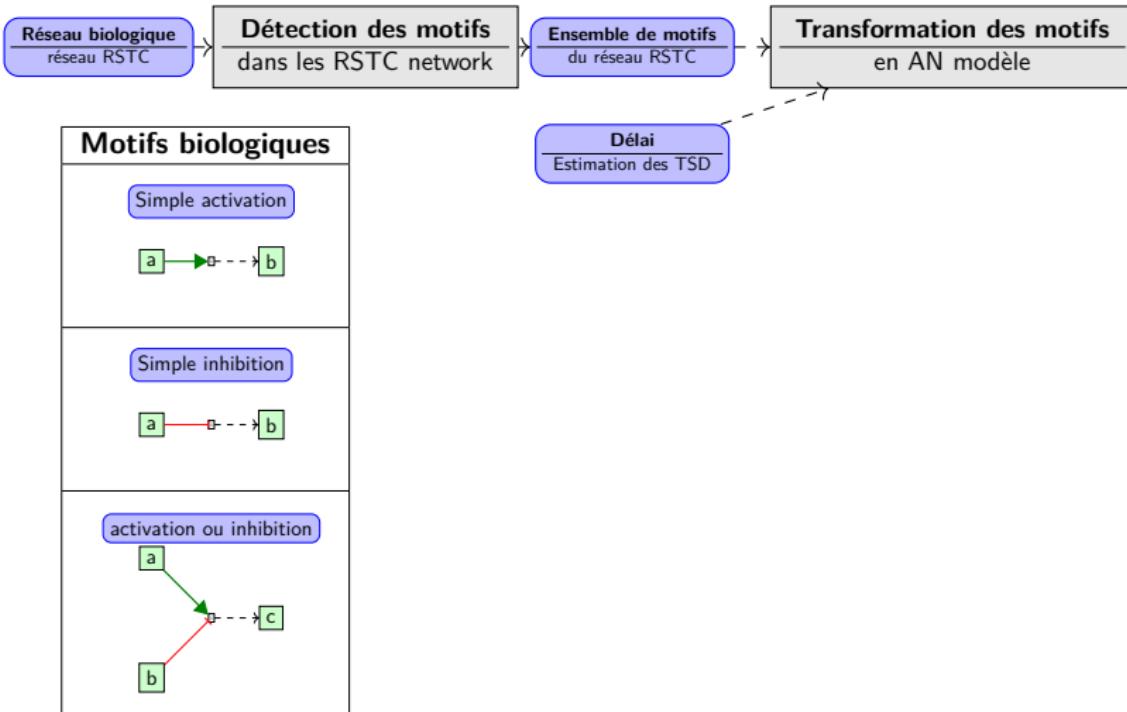
Formalisation des réseaux RSTC

Détection et transformation automatique des motifs minimaux



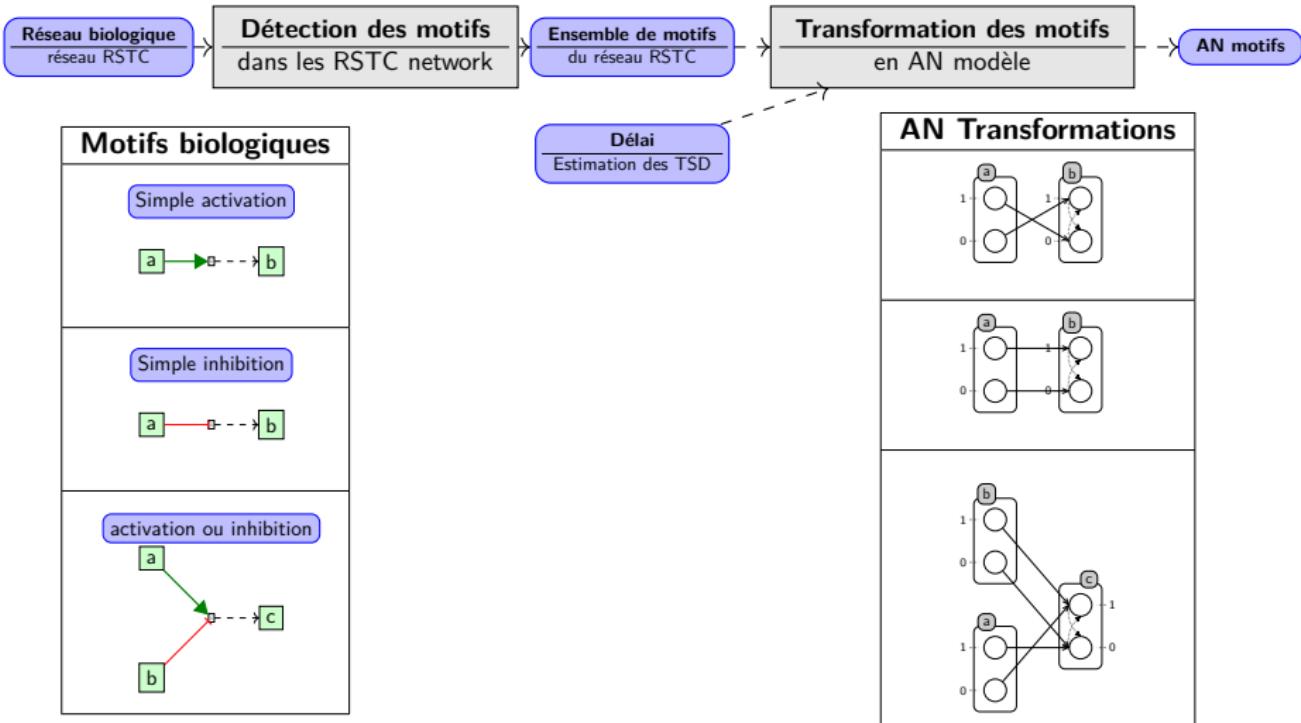
Formalisation des réseaux RSTC

Détection et transformation automatique des motifs minimaux



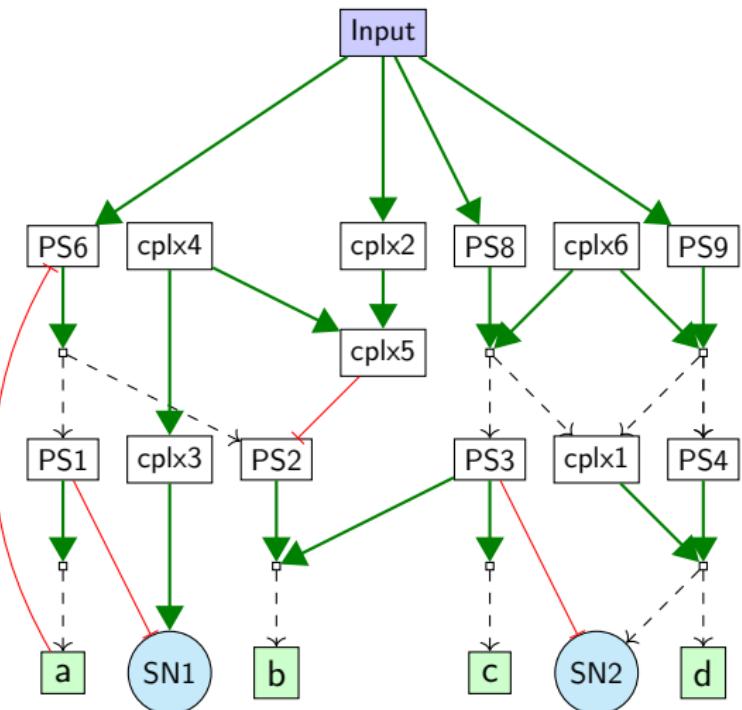
Formalisation des réseaux RSTC

Détection et transformation automatique des motifs minimaux



Formalisation des réseaux RSTC

Détection automatique des motifs minimaux



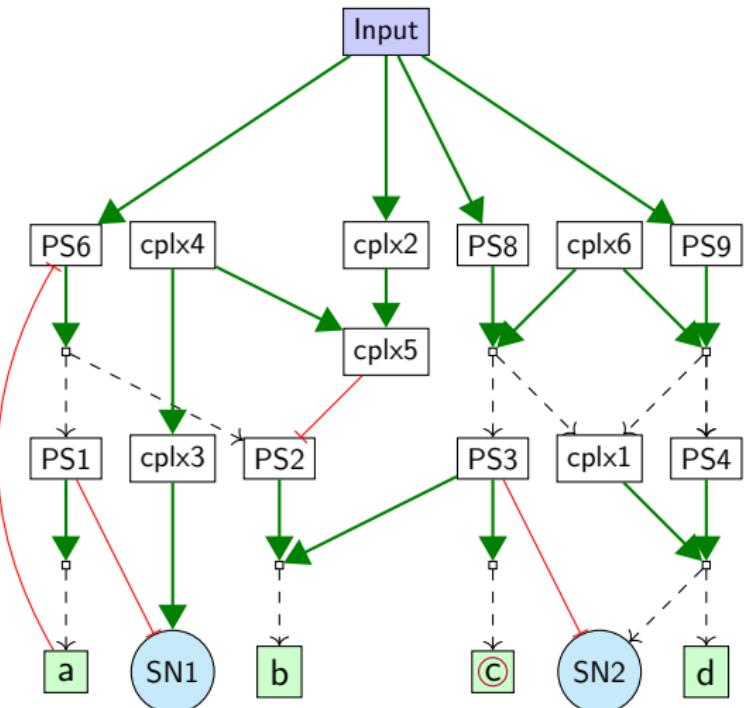
Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

	Complexe, protéine
	gène
	état cellulaire
▫	nœud transitoire

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux



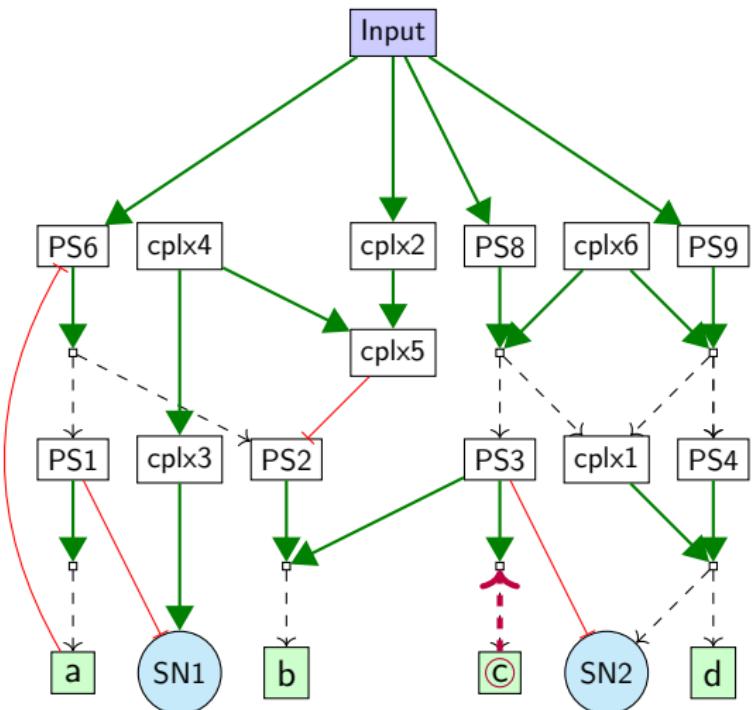
Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

	Complexe, protéine
	gène
	état cellulaire
▫	nœud transitoire

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux

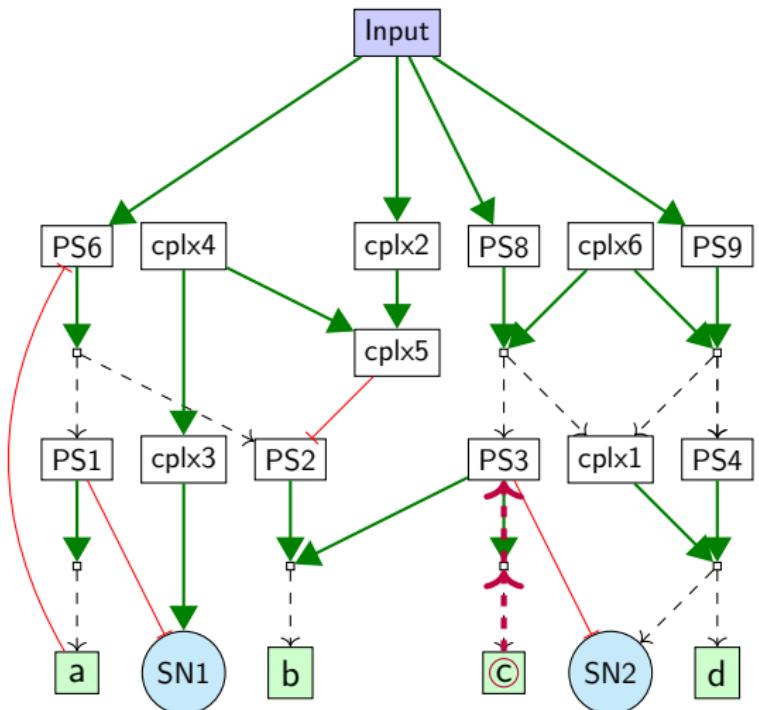


Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux



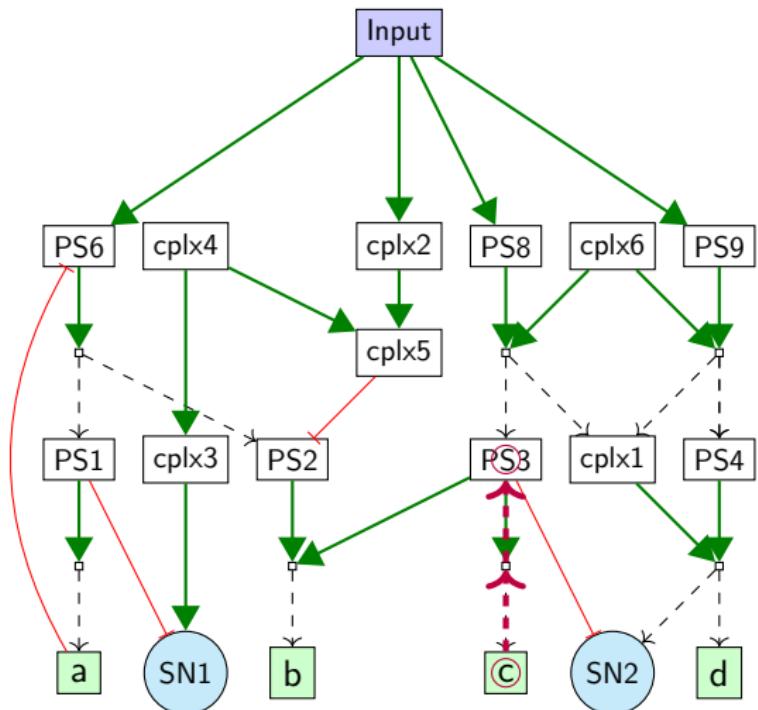
Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

	Complexe, protéine
	gène
	état cellulaire
▫	nœud transitoire

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux



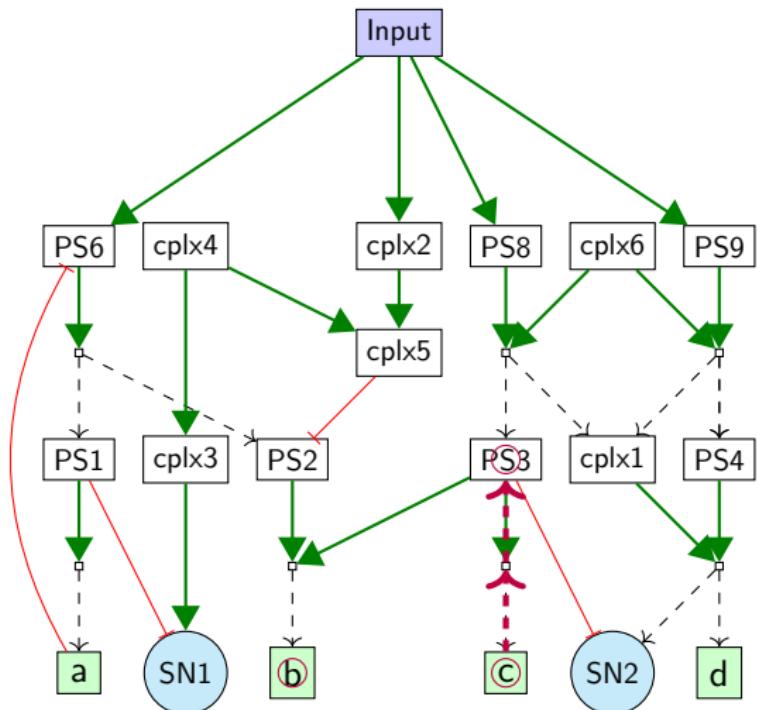
Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

	Complexe, protéine
	gène
	état cellulaire
▫	nœud transitoire

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux



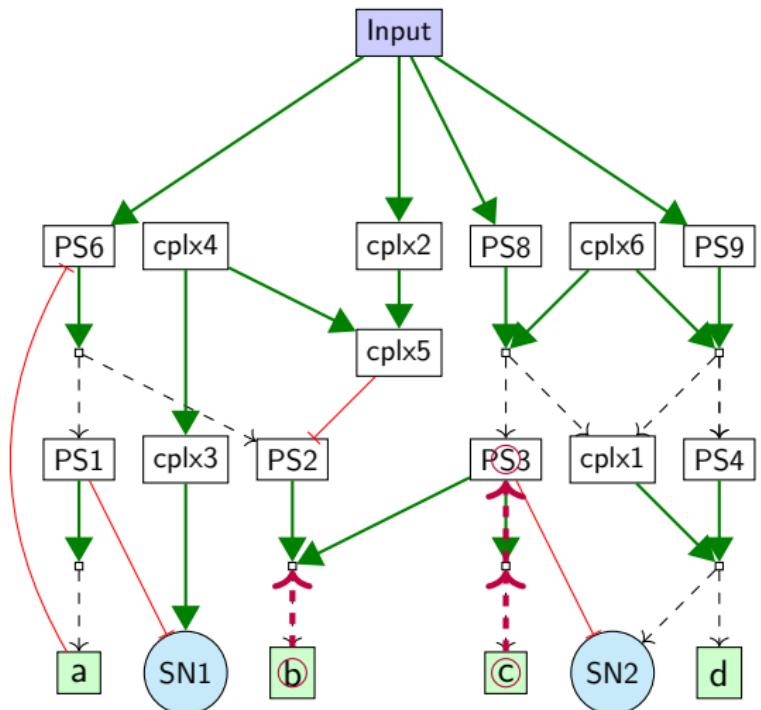
Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

	Complexe, protéine
	gène
	état cellulaire
▫	nœud transitoire

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux



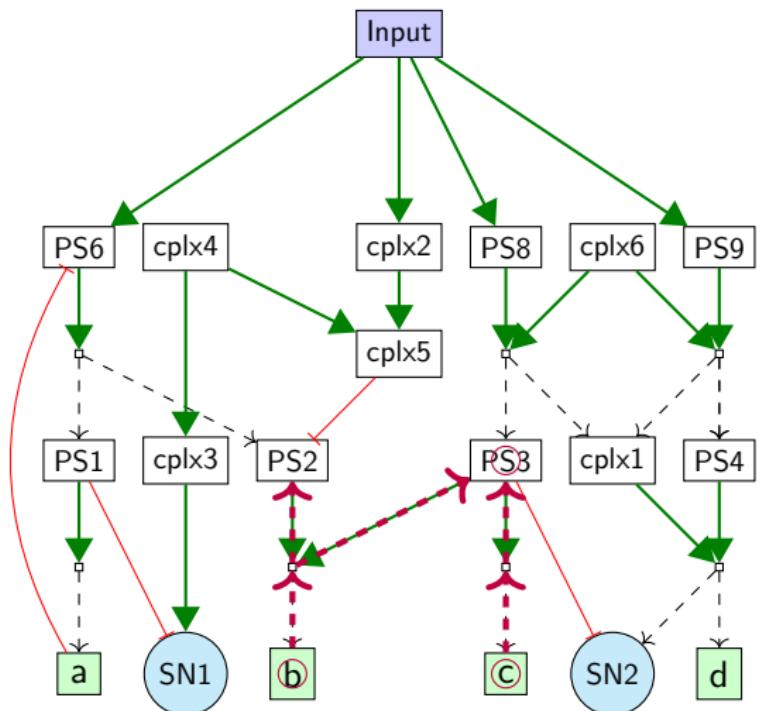
Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

	Complexe, protéine
	gène
	état cellulaire
▫	nœud transitoire

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux

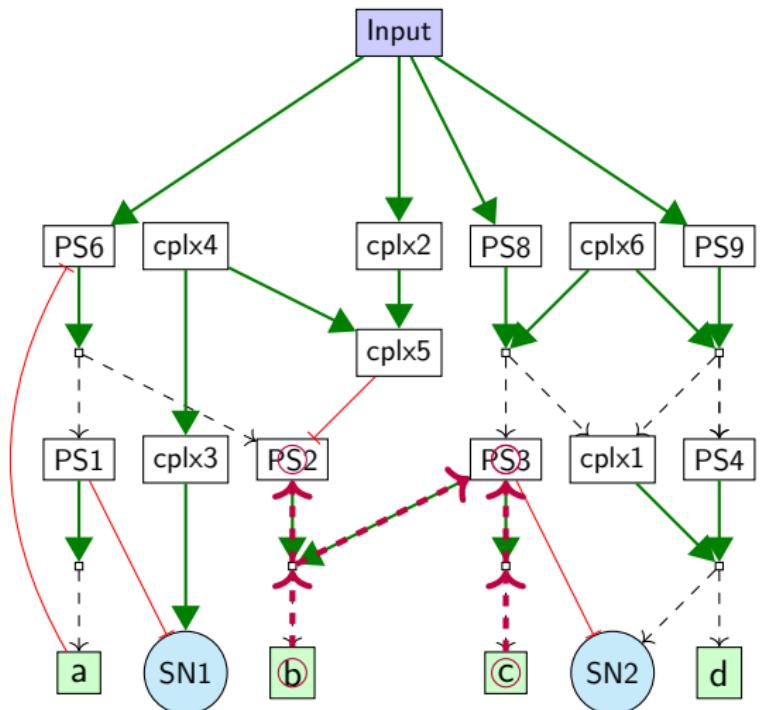


Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux

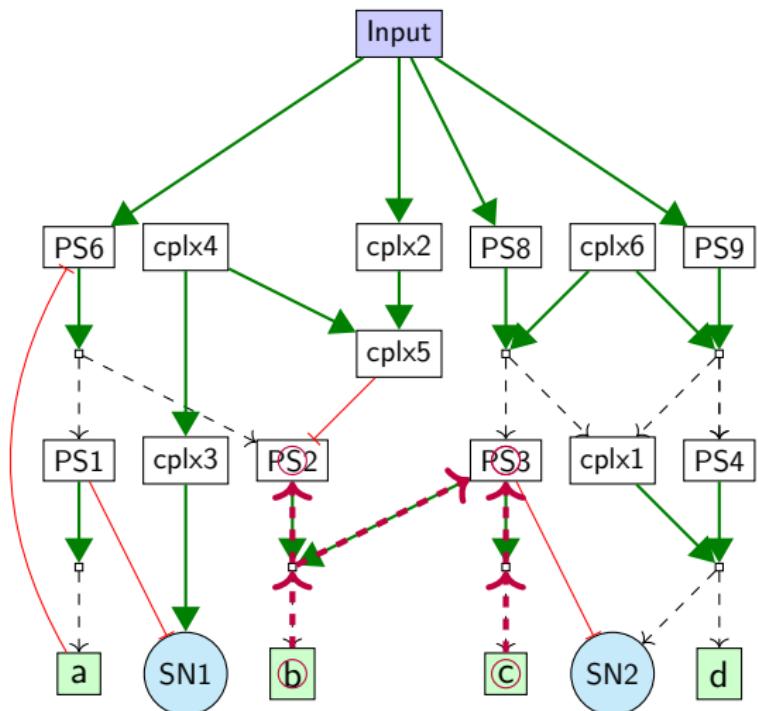


Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires:
translocations, modifications, transcriptions,...

Formalisation des réseaux RSTC

Détection automatique des motifs minimaux



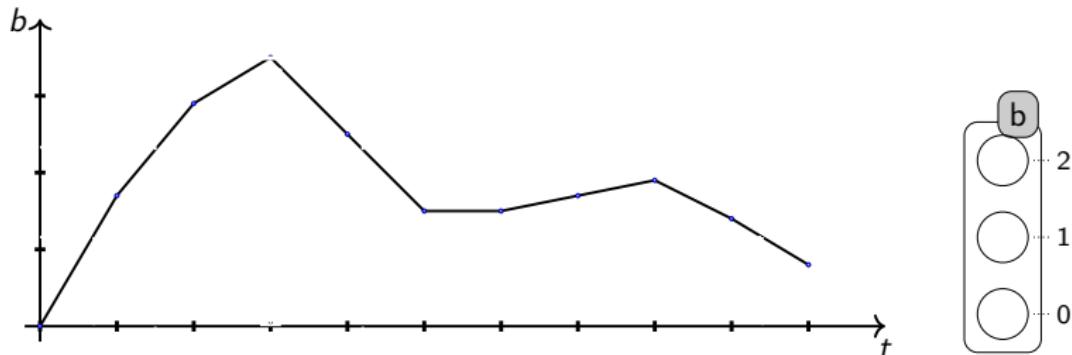
Nœuds terminaux: protéines, complexes, gènes, états cellulaires,...

Nœuds transitoires: translocations, modifications, transcriptions,...

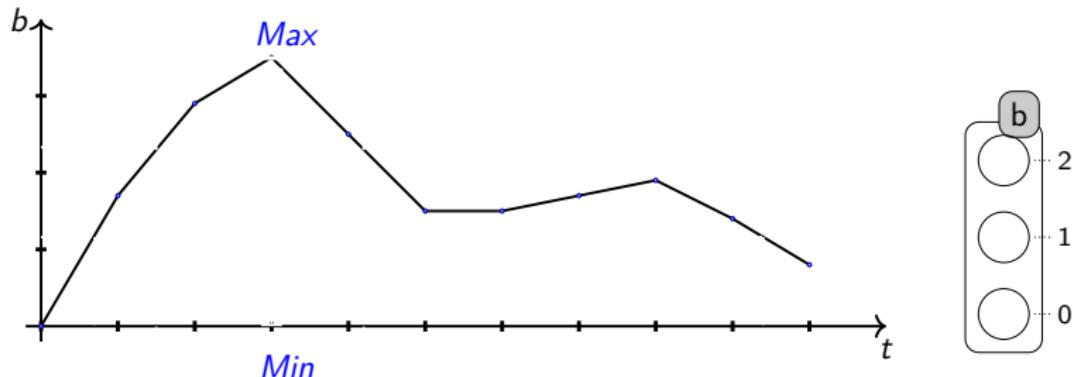
Complexité: $\mathcal{O}(|V| \log(h))$. avec h = hauteur moyenne des motifs.

	Complexe, protéine
	gène
	état cellulaire
▫	nœud transitoire

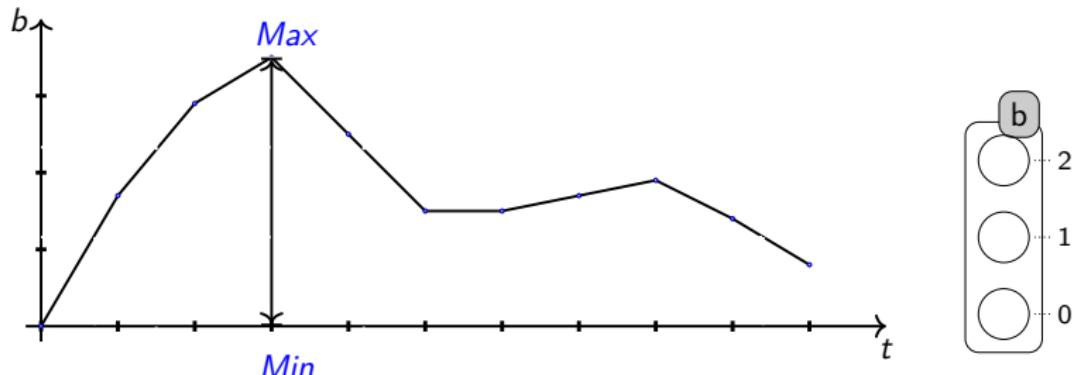
Estimation du taux des séries temporelles



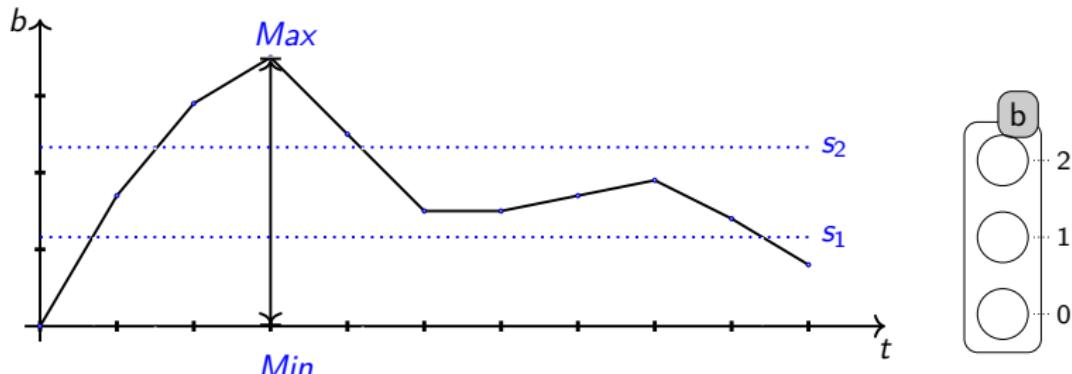
Estimation du taux des séries temporelles



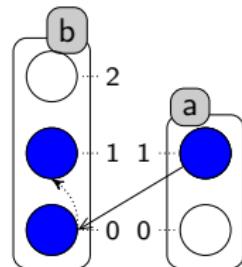
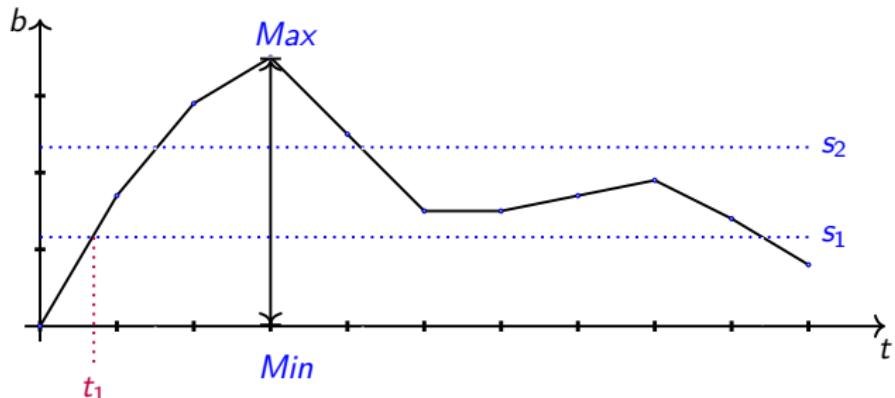
Estimation du taux des séries temporelles



Estimation du taux des séries temporelles

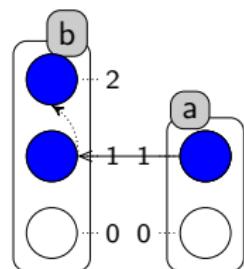
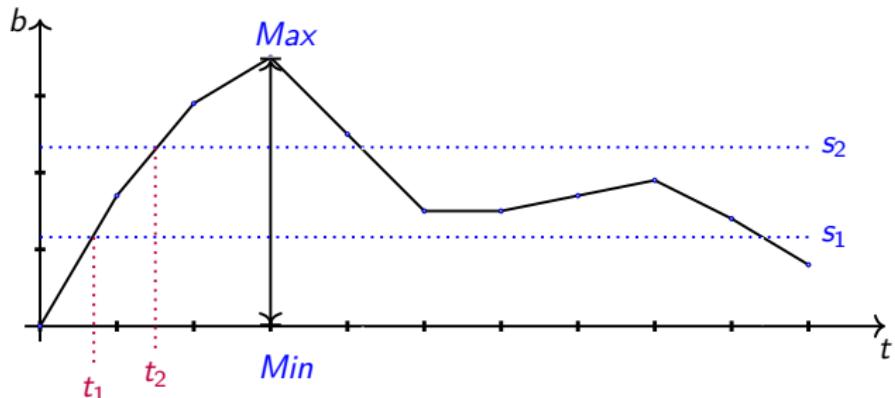


Estimation du taux des séries temporelles



$$a_1 \rightarrow b_0 \uparrow b_1 \text{ avec } r_1 = \frac{1}{t_1 - t_0}$$

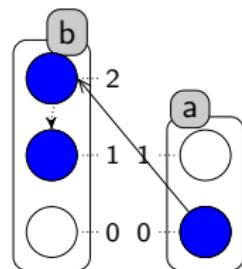
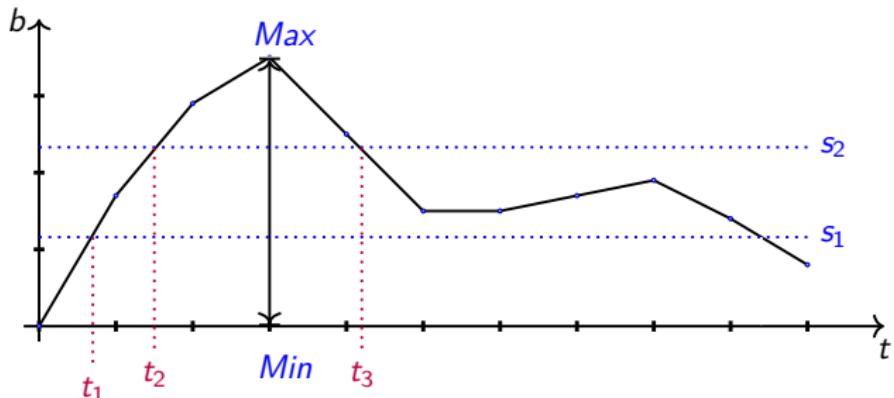
Estimation du taux des séries temporelles



$$a_1 \rightarrow b_0 \uparrow b_1 \text{ avec } r_1 = \frac{1}{t_1 - t_0}$$

$$a_1 \rightarrow b_1 \uparrow b_2 \text{ avec } r_2 = \frac{1}{t_2 - t_1}$$

Estimation du taux des séries temporelles

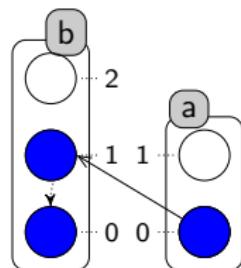
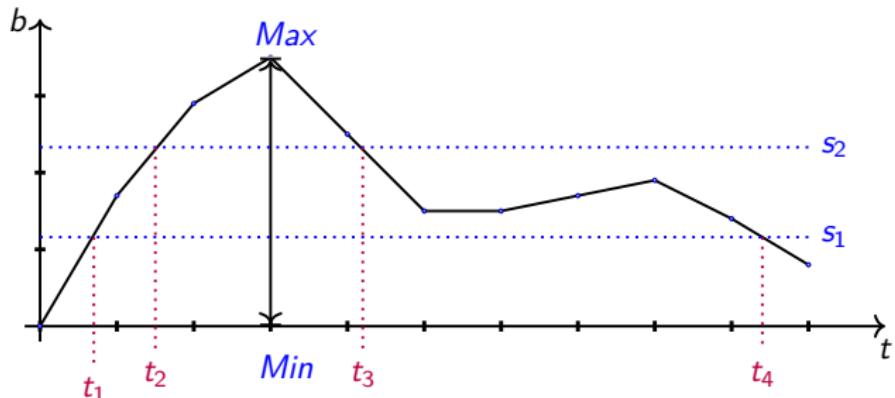


$$a_1 \rightarrow b_0 \xrightarrow{r} b_1 \text{ avec } r_1 = \frac{1}{t_1 - t_0}$$

$$a_1 \rightarrow b_1 \xrightarrow{r} b_2 \text{ avec } r_2 = \frac{1}{t_2 - t_1}$$

$$a_0 \rightarrow b_2 \xrightarrow{r} b_1 \text{ avec } r_3 = \frac{1}{t_3 - t_2}$$

Estimation du taux des séries temporelles



$$a_1 \rightarrow b_0 \uparrow b_1 \text{ avec } r_1 = \frac{1}{t_1 - t_0}$$

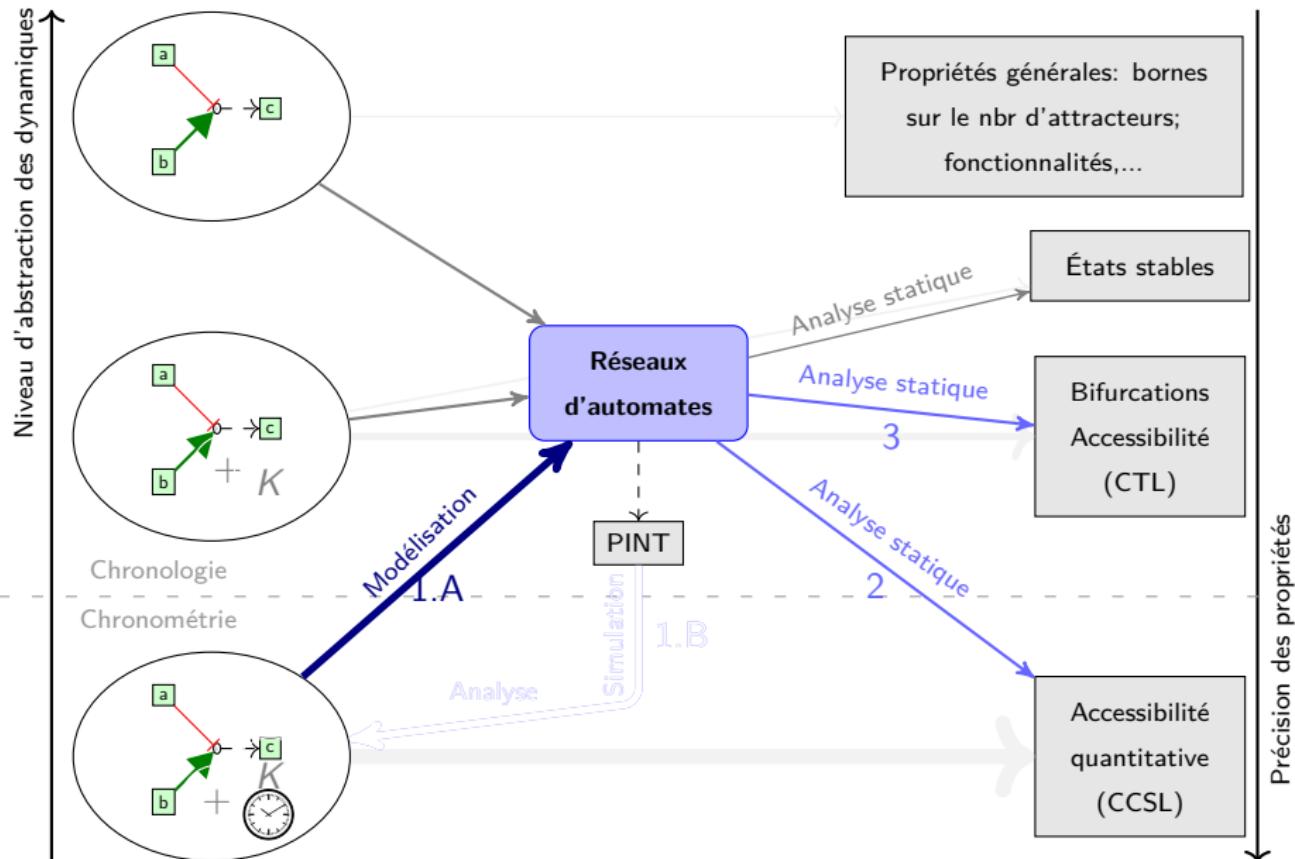
$$a_1 \rightarrow b_1 \uparrow b_2 \text{ avec } r_2 = \frac{1}{t_2 - t_1}$$

$$a_0 \rightarrow b_2 \uparrow b_1 \text{ avec } r_3 = \frac{1}{t_3 - t_2}$$

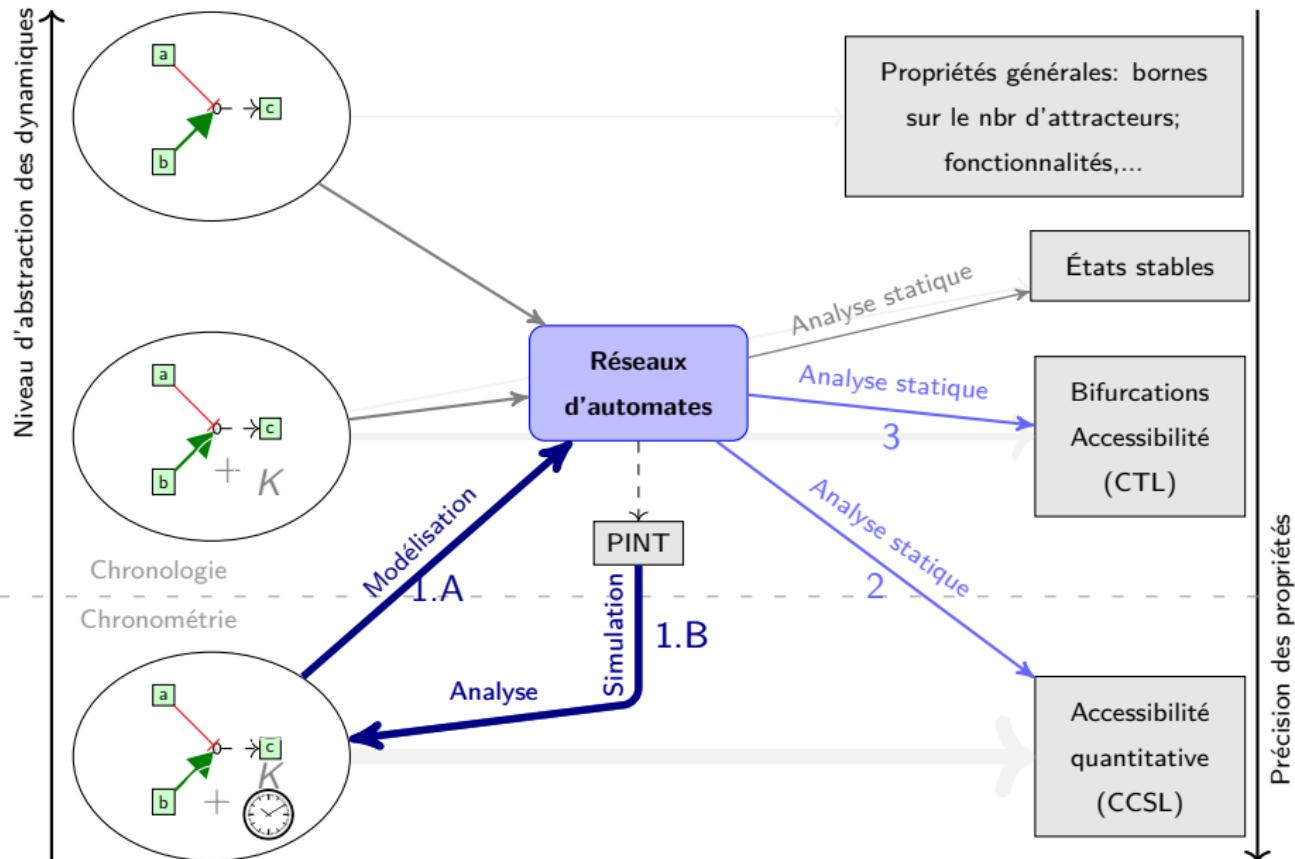
$$a_0 \rightarrow b_1 \uparrow b_0 \text{ avec } r_4 = \frac{1}{t_4 - t_3}$$

$$\text{Généralisation: } r_i = \frac{1}{t_i - t_{i-1}}$$

Contributions

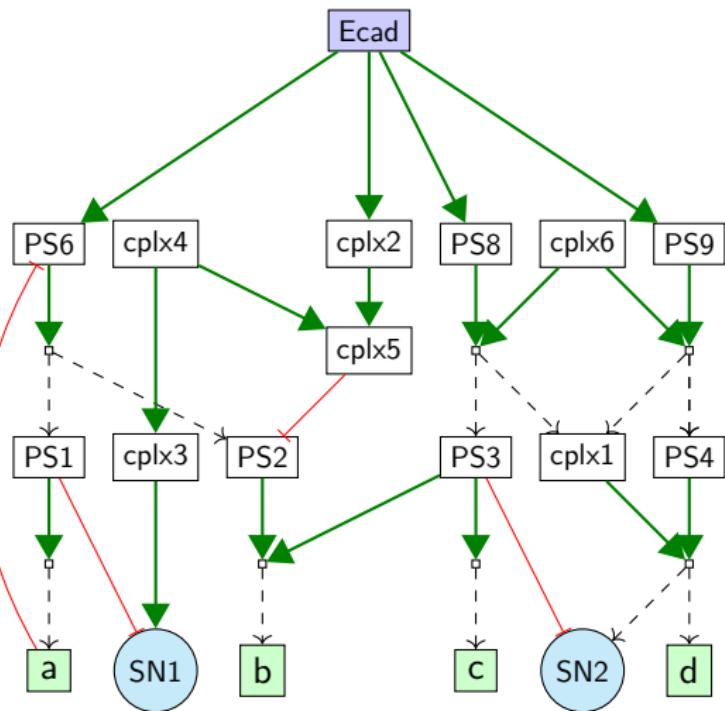


Contributions



Hypothèses de simulation

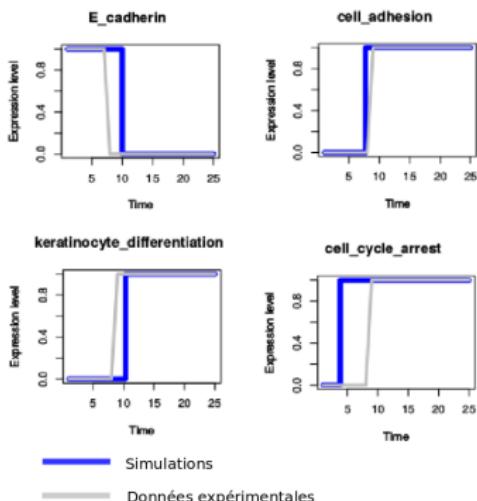
Hypothèses structurelles et dynamiques



- Introduction d'un signal d'entrée: Ecad.
- Introduction des **auto-transitions** pour les facteurs de transcription.
- Supposons **trois niveaux d'expression** pour les gènes.

Simulations et analyses

Simulation des processus cellulaire



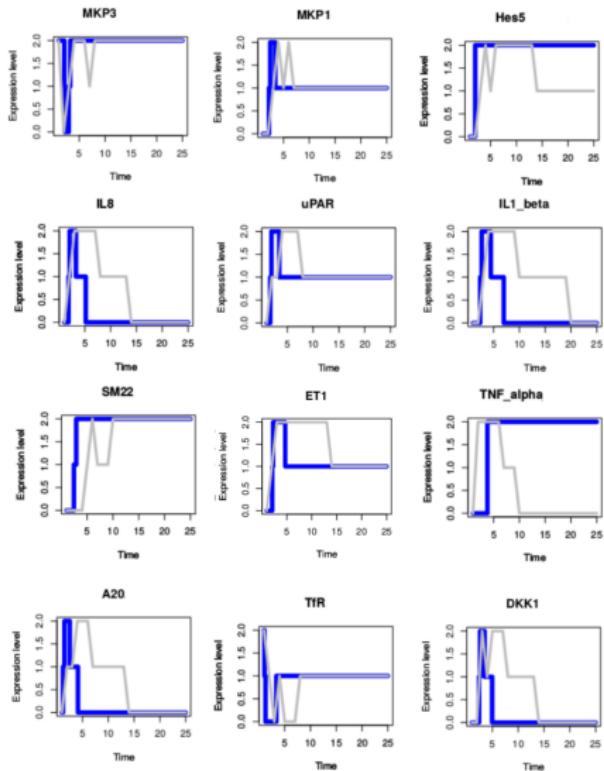
Simulations

- Ecad conforme avec l'expérience et la modélisation.
- Dynamique des processus cellulaire conforme avec la littérature.

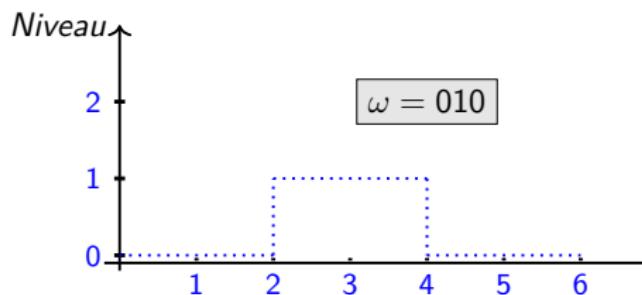
[Kolly et al, in *J. Invest. Dermatol.*, 2005], [Tu et al, in *J. Invest. Dermatol.*, 2011]

Simulations et analyses

Simulations des gènes



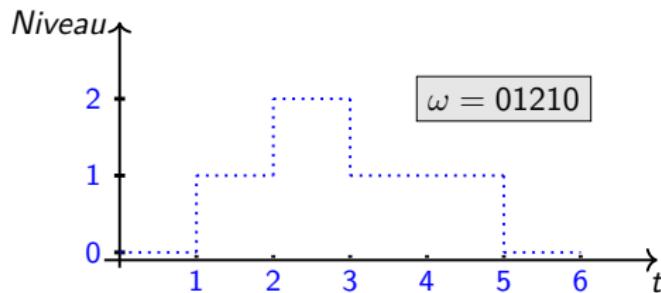
Simulation et analyse statistique des traces des simulations



Pour chaque composant C_i , $1 \leq i \leq P$, N simulations génèrent $\omega_{i1}, \omega_{i2}, \dots, \omega_{iN}$ mots.

Pour $1 \leq j \leq N$

$\omega_{ij} \Rightarrow$ \mathcal{A}_{C_i} \Rightarrow oui / non



$$\% \text{Validation} = \frac{|\text{oui}|}{|\text{Simulations}|}$$

Simulation et analyse statistique des traces des simulations

Automate	composant	% validation	% validation T_1
$\mathcal{A}_2(01210)$	A20	91	100
$\mathcal{A}_2(01210)$	IL1_beta	81	100
$\mathcal{A}_2(01210)$	IL8	93	100
$\mathcal{A}_2(01210)$	TNF_alpha	0	0
$\mathcal{A}_3(0121)$	uPar	76	99
$\mathcal{A}_3(0121)$	ET1	8	19
$\mathcal{A}_4(0121210)$	DKK1	13	43
$\mathcal{A}_5(012121)$	Hes5	0	17
$\mathcal{A}_5(012121)$	MKP1	9	97
$\mathcal{A}_6(0212)$	SM22	11	100
$\mathcal{A}_7(02010)$	MKP3	11	98
$\mathcal{A}_8(02121)$	Tfr	0	94

Nombre de simulation: 100

T_1 : Tolérance de type T_1 .

- Raffinement de la dynamique par intégration des séries temporelles ◦ Analyse statistique des traces

Synthèse sur l'intégration des données

Modélisation hybride des réseaux RSTC

Formalisation des connaissances biologiques

- Identification des motifs minimaux dans les RSTC.
- Raffinement structurel de la dynamique avec des automates de synchronisation.

Intégration des paramètres temporels et stochastiques & simulation

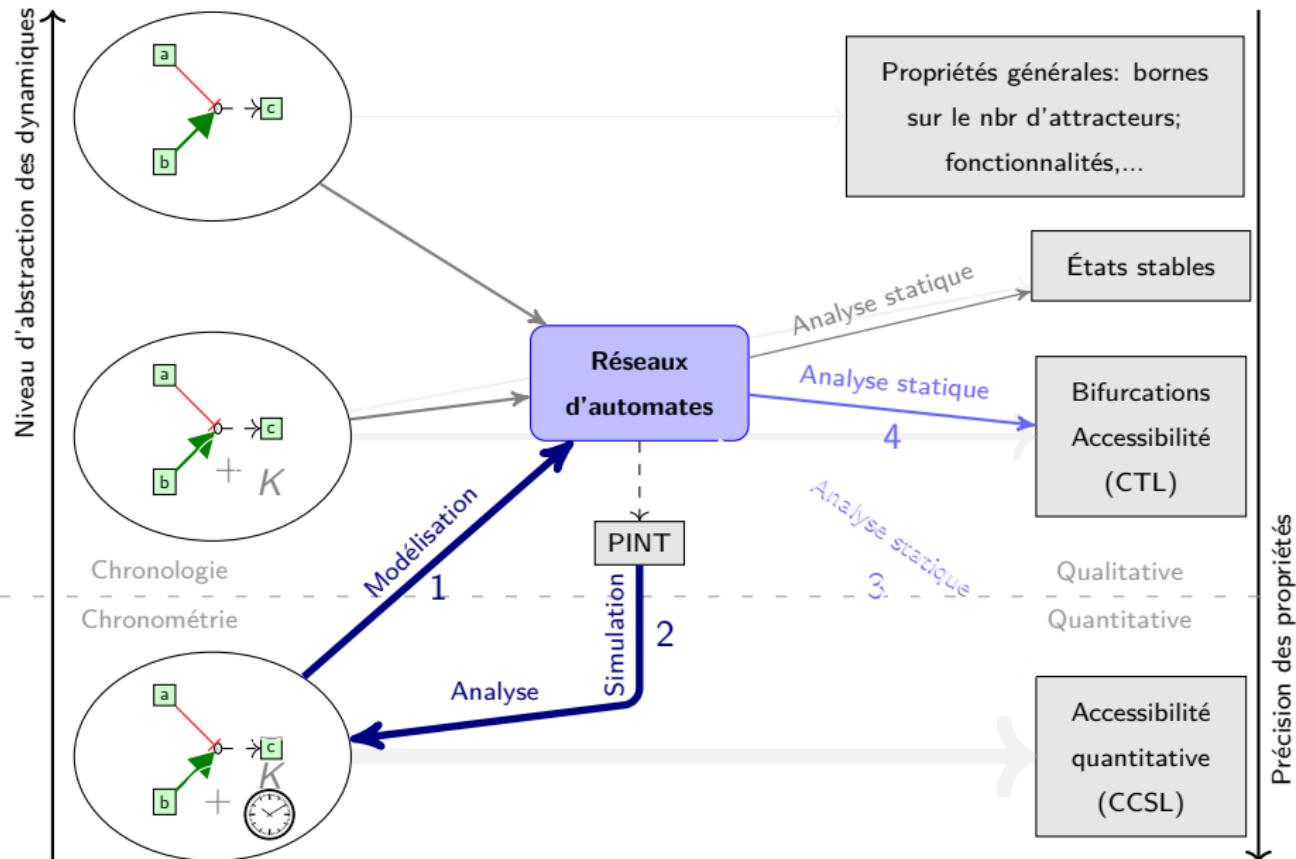
- Inférence des paramètres des données de séries temporelles.
- Simulation stochastique sur des modèles biologiques réels.
- Analyse statistique des traces des simulations.

Publication

- Louis Fippo Fitime, Christian Schuster, Peter Angel, Olivier Roux, Carito Guziolowski.
Integrating time-series data in large-scale discrete cell-based models. In A. Abate and D. Safranek, editors, *4th International Workshop on Hybrid Systems Biology (HSB'2015)*, Lecture Notes in Computer Science. September 2015.

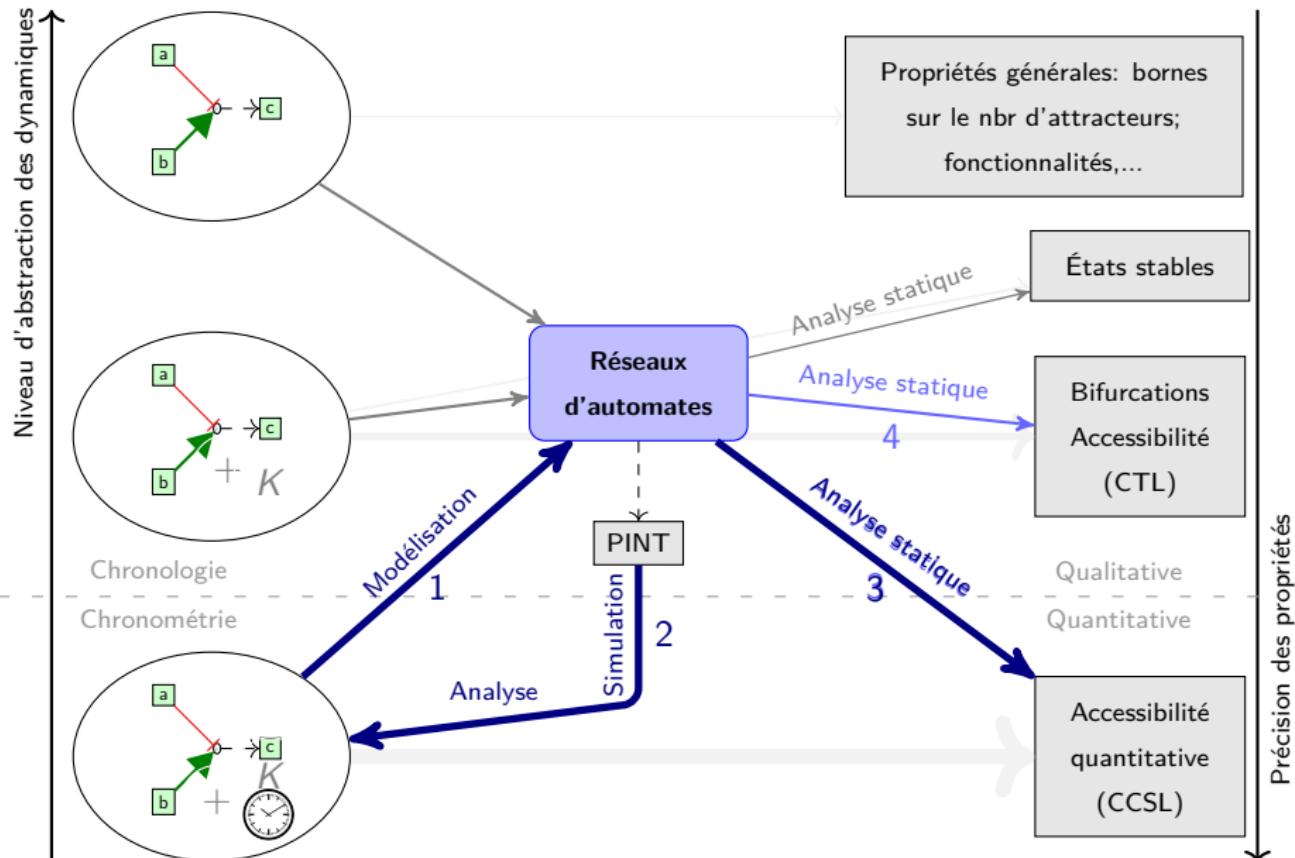
- Analyse statique des propriétés quantitatives dans les réseaux d'automates stochastiques

Contributions



- Analyse statique des propriétés quantitatives dans les réseaux d'automates stochastiques

Contributions

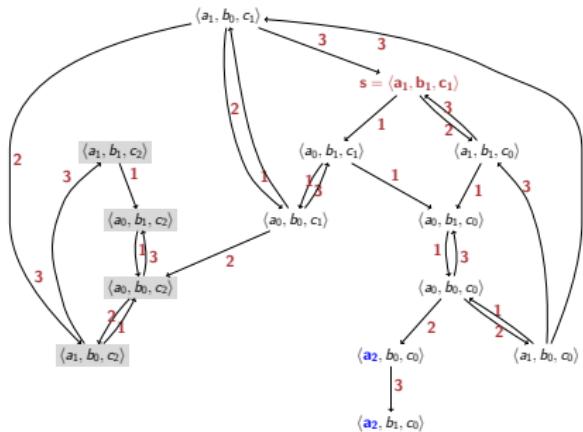


- Analyse statique des propriétés quantitatives dans les réseaux d'automates stochastiques

SAN abstraits comme CTMC

La probabilité d'effectuer une transition depuis l'état s :

$$\mathcal{P}_{(s,*)}(t) = 1 - e^{-\lambda t}$$



Inconvénient: **explosion de l'espace d'états.**

Probabilité d'effectuer une transition de s à s' :

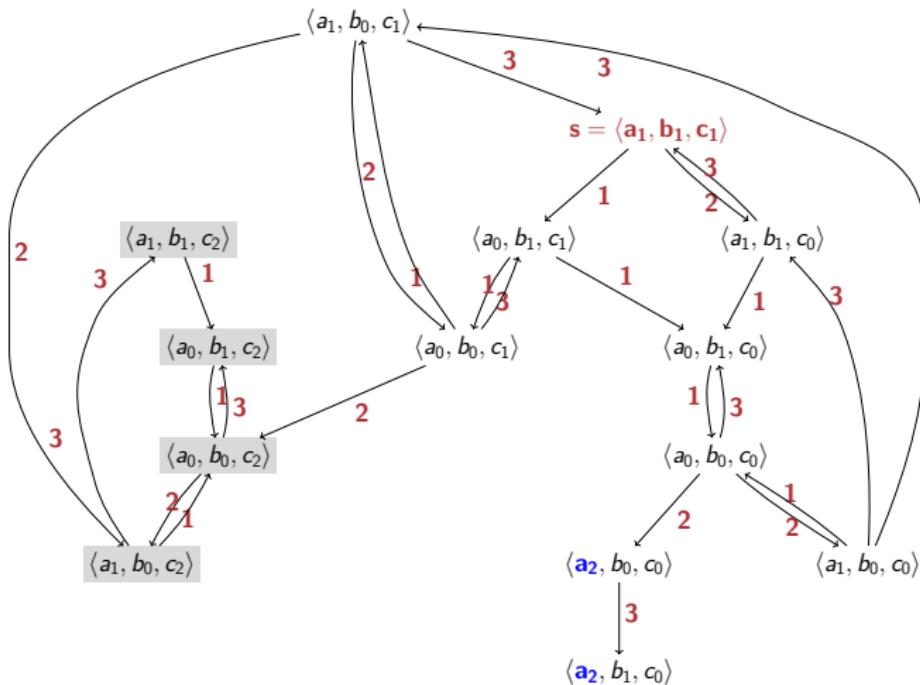
$$\mathcal{P}_{(s,s')}(t) = \frac{R(s,s')}{E(s)} \cdot \mathcal{P}_{(s,*)}(t)$$

Probabilité d'effectuer une transition de s à un ensemble d'états:

$$\mathcal{P}_{(s,A)}(t) = \frac{R(s,A)}{E(s)} \cdot \mathcal{P}_{(s,*)}(t)$$

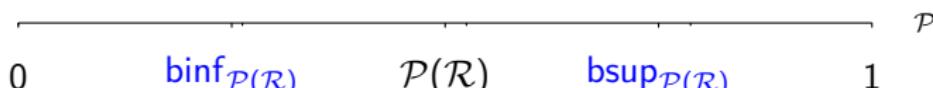
- Analyse statique des propriétés quantitatives dans les réseaux d'automates stochastiques

SAN abstraits comme CTMC



Inconvénient: explosion de l'espace d'états.

Contribution



But:

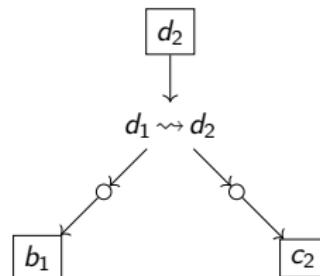
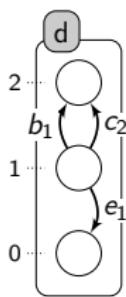
- Analyse statique des propriétés quantitatives.
- Estimer une borne inférieure et supérieure de la probabilité et du délai.

Contribution:

$$\left\{ \begin{array}{l} \text{QS (Classes d'équivalences d'états)} \\ + \\ \text{IA (Interprétation abstraite)} \end{array} \right. \implies \text{approximation formelle des propriétés quantitatives}$$

Interprétation abstraite

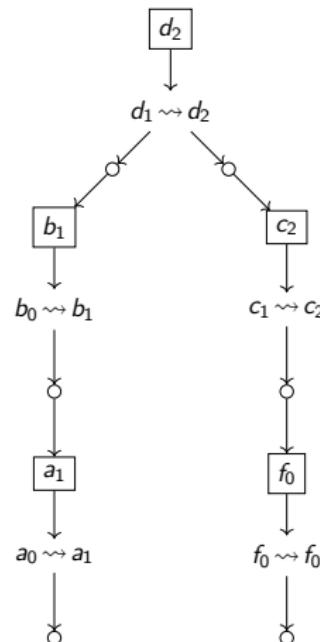
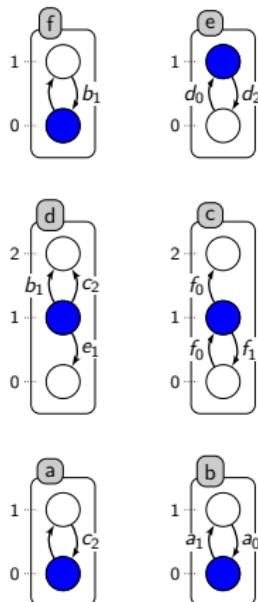
Causalité locale



Interprétation abstraite

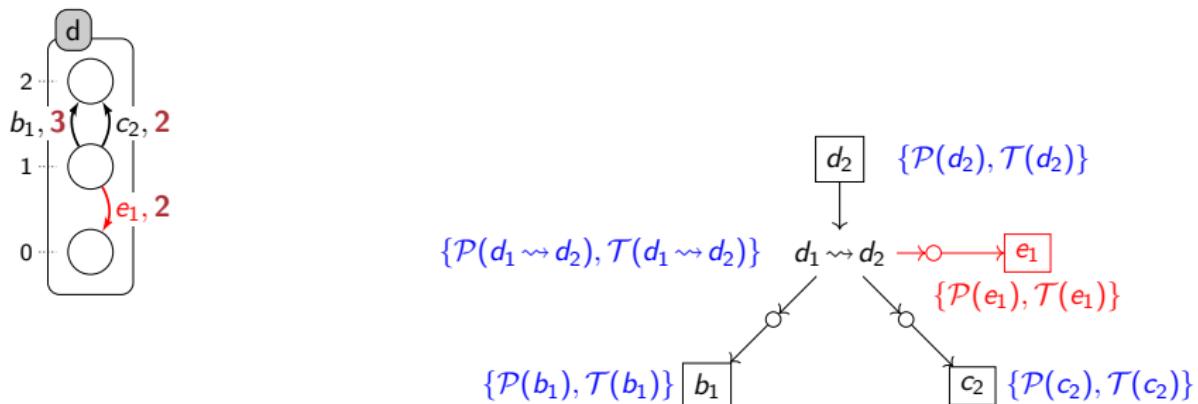
Graphe de causalité locale

- **Contexte initial/état:** $\langle a_0, b_0, c_1, d_1, e_1, f_0 \rangle$
- **Goal:** état local d_2



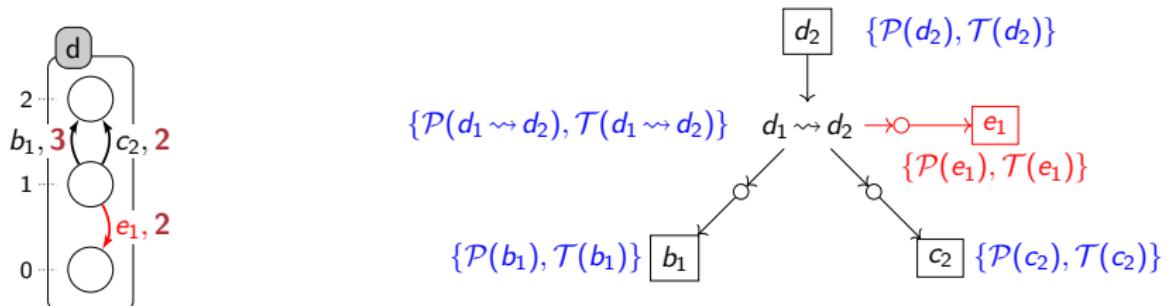
Interprétation abstraite

Causalité locale quantifiée & Classes d'équivalence d'états



Interprétation abstraite

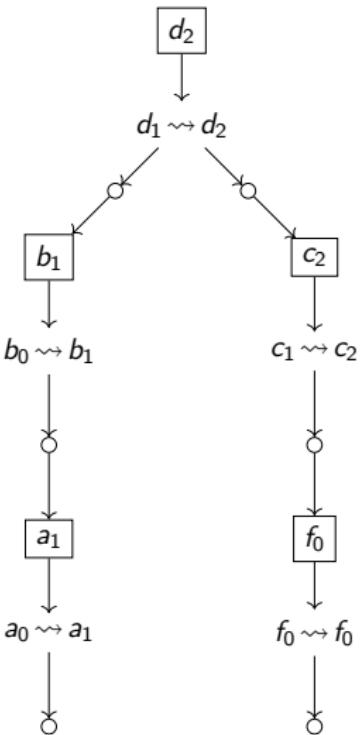
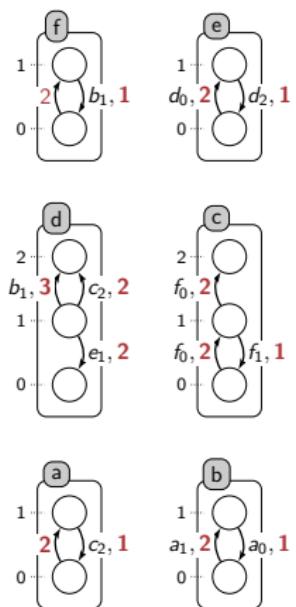
Causalité locale quantifiée & Classe d'équivalences d'états



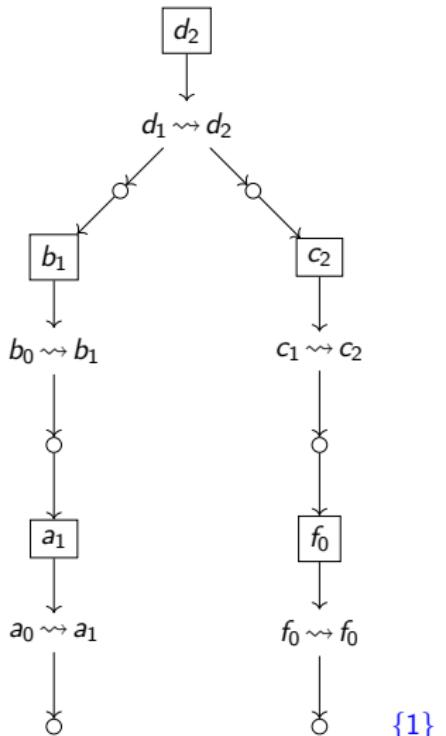
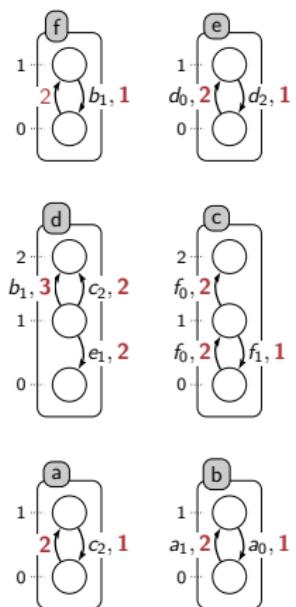
$$\mathcal{P}_t(d_1 \rightsquigarrow d_2)(t) = \frac{3 \cdot \mathcal{P}(b_1) + 2 \cdot \mathcal{P}(c_2)}{3 \cdot \mathcal{P}(b_1) + 2 \cdot \mathcal{P}(c_2) + 2 \cdot \mathcal{P}(e_1)} \cdot \mathcal{P}_{(A_{d_1}, *)}(t)$$

Approximation: $\mathcal{P}(e_1) = 1$

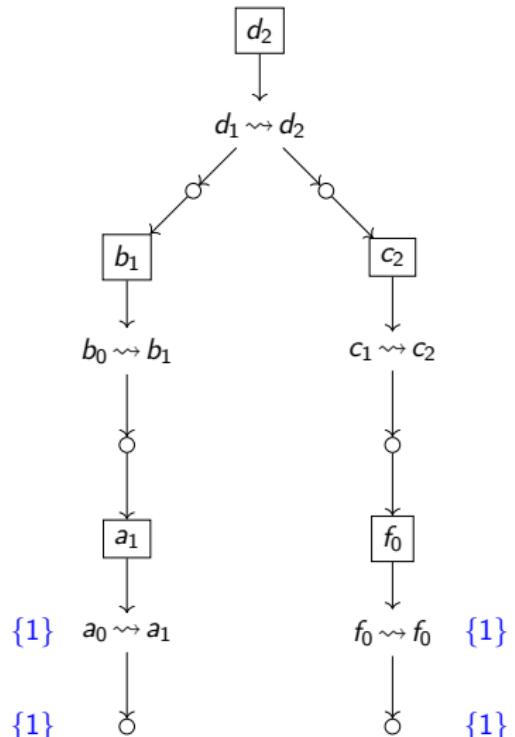
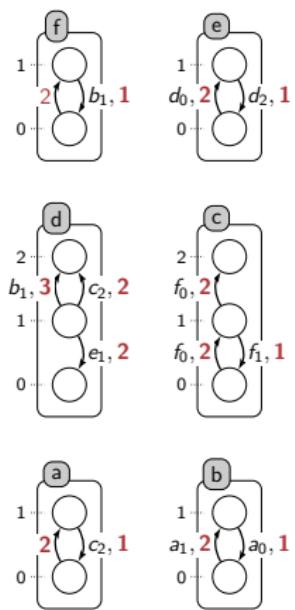
Approximation de la probabilité d'accessibilité



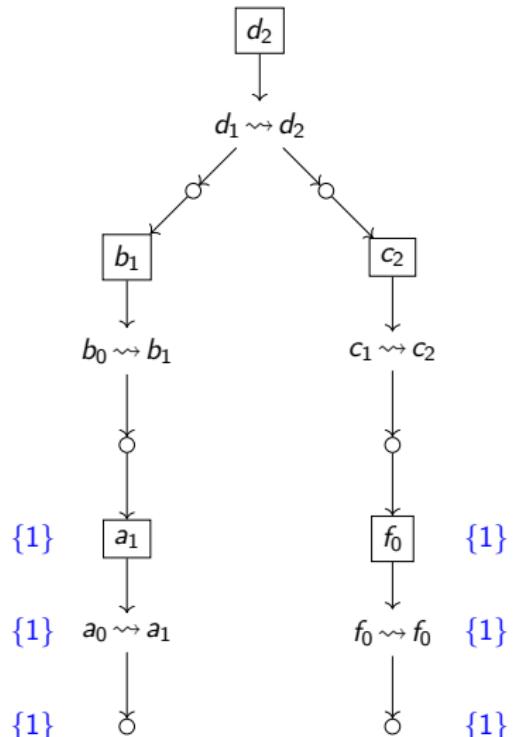
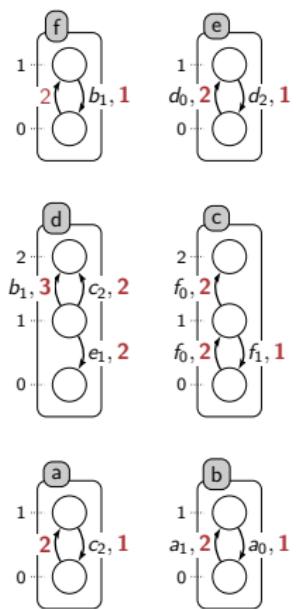
Approximation de la probabilité d'accessibilité



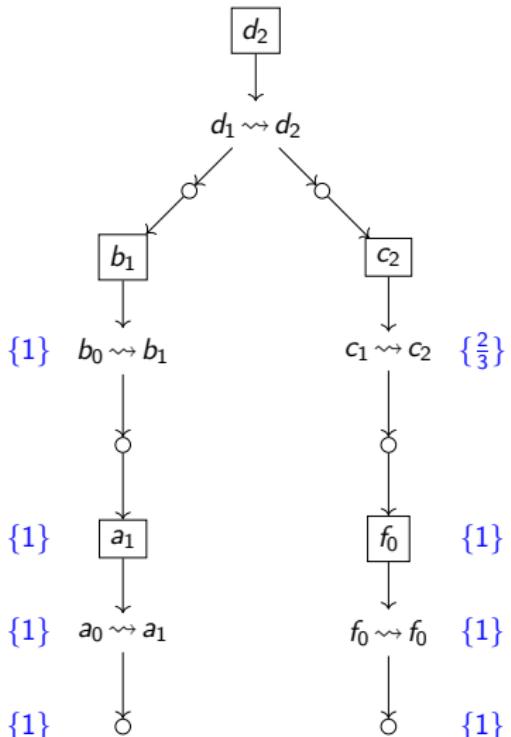
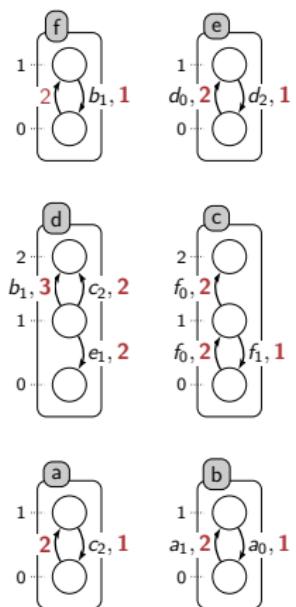
Approximation de la probabilité d'accessibilité



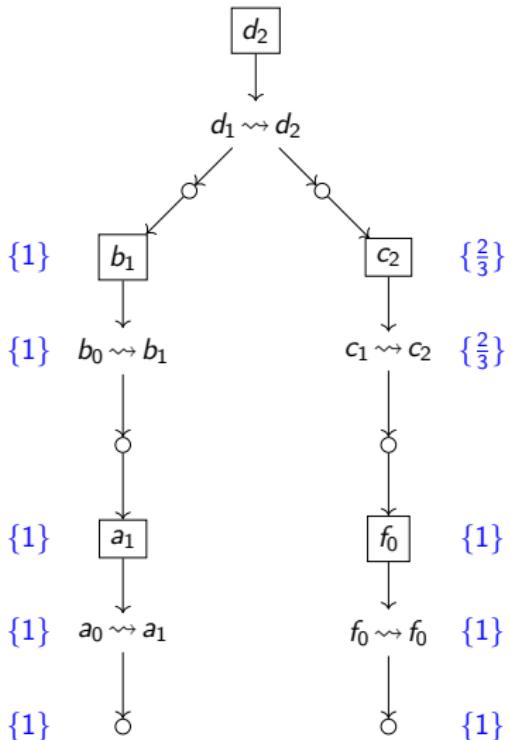
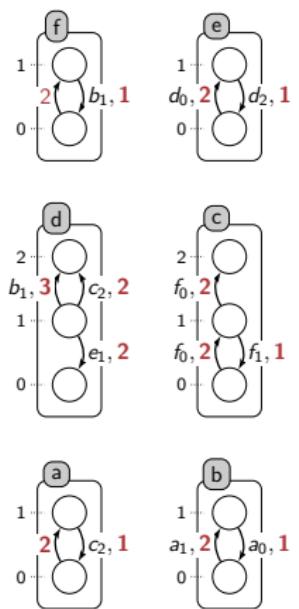
Approximation de la probabilité d'accessibilité



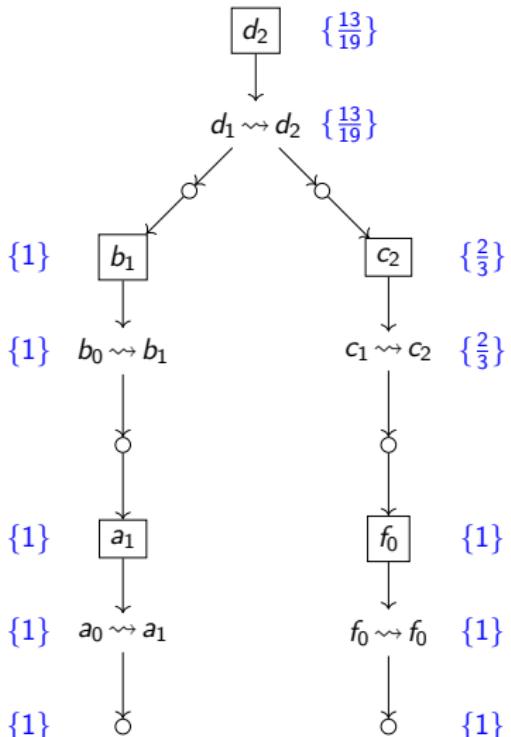
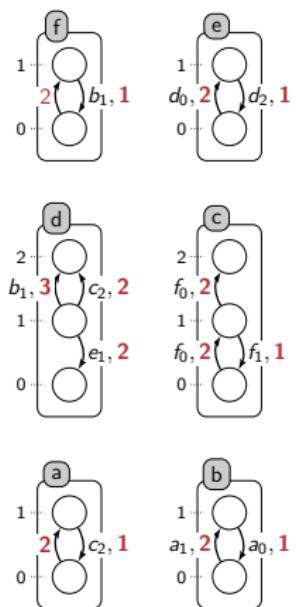
Approximation de la probabilité d'accessibilité



Approximation de la probabilité d'accessibilité



Approximation de la probabilité d'accessibilité



Synthèse sur l'analyse statique des propriétés quantitatives

Approximation des propriétés quantitatives

Interprétation abstraite des scénarios

- Causalité locale quantitative.
- Graphe de causalité locale quantifié.

Outil d'aide à la décision

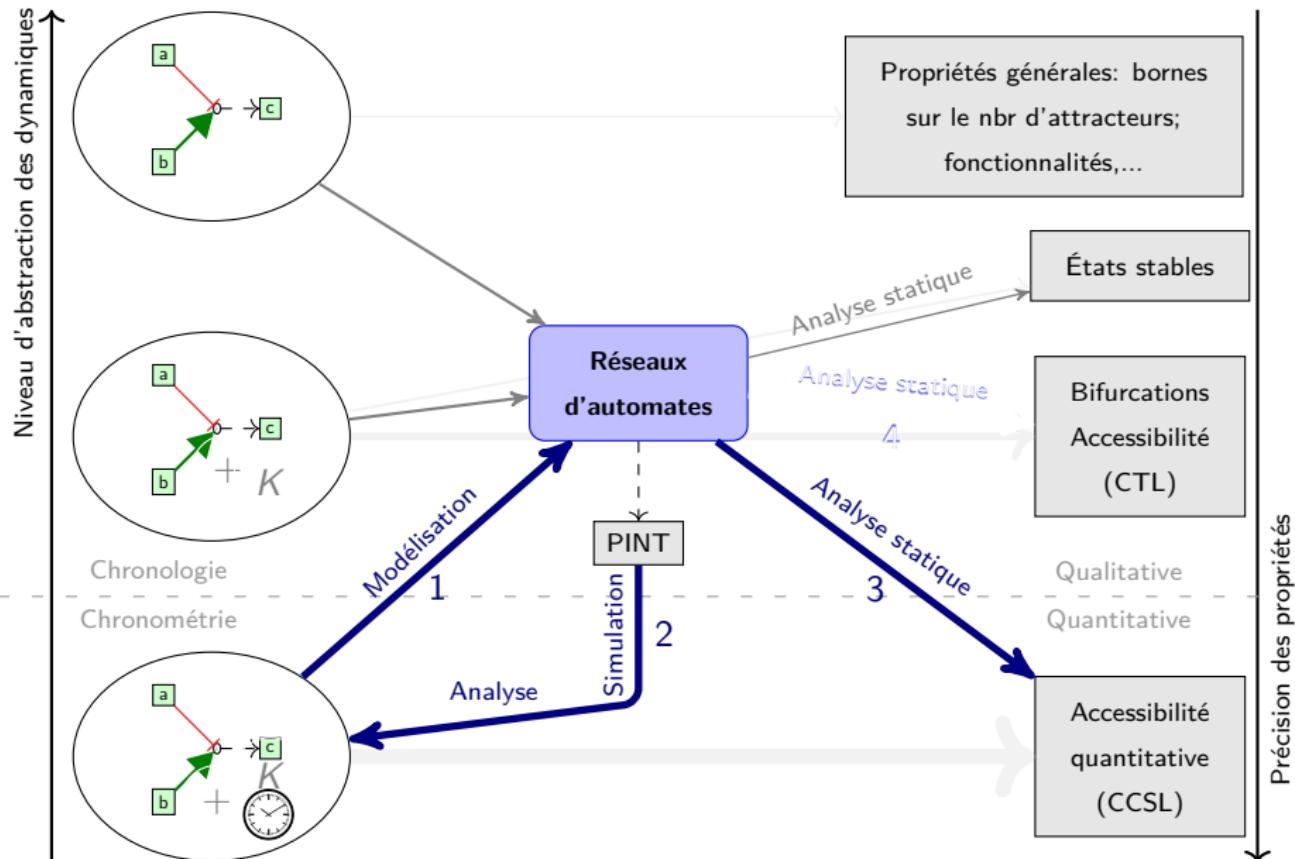
- Borne inférieure de la probabilité et des délais.
- Choix des scénarios (plus probable, plus rapide,...).

Intérêt:

Exprimer les formules comme: quelle est la probabilité d'observer une protéine donnée a après avoir observé une protéine b ?

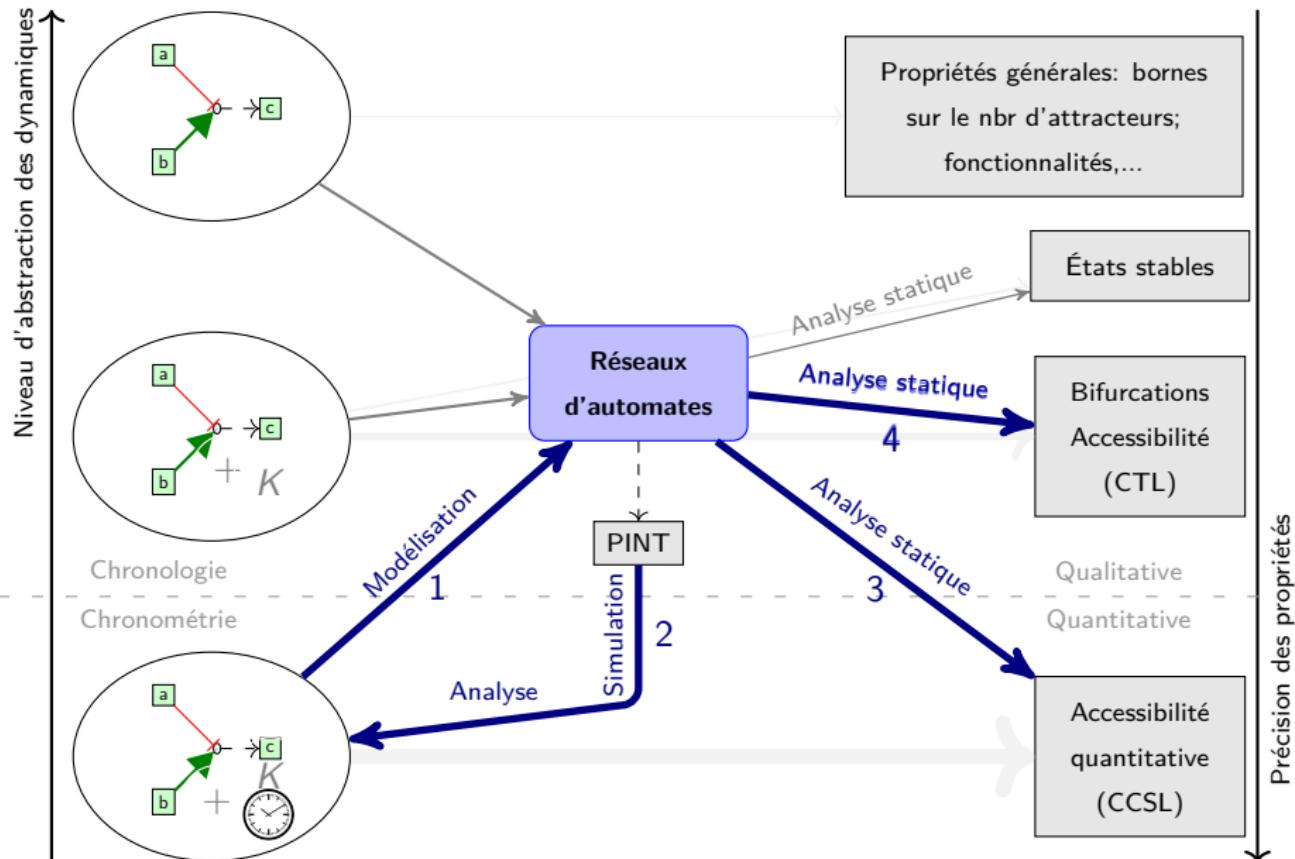
- Analyse statique des bifurcations

Contributions



- Analyse statique des bifurcations

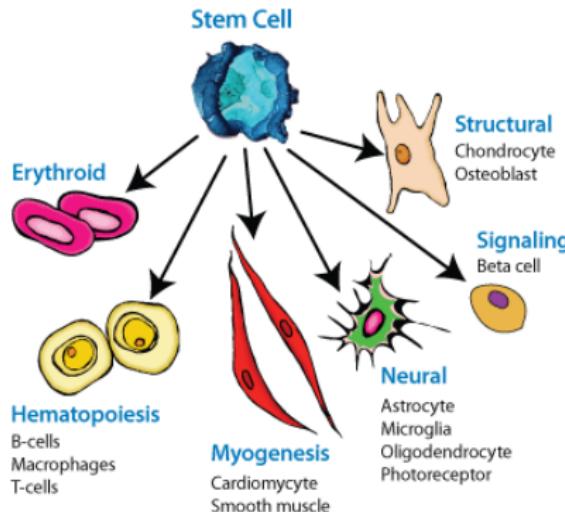
Contributions



- Analyse statique des bifurcations

Introduction à la notion de bifurcation

Différenciation des cellules souches

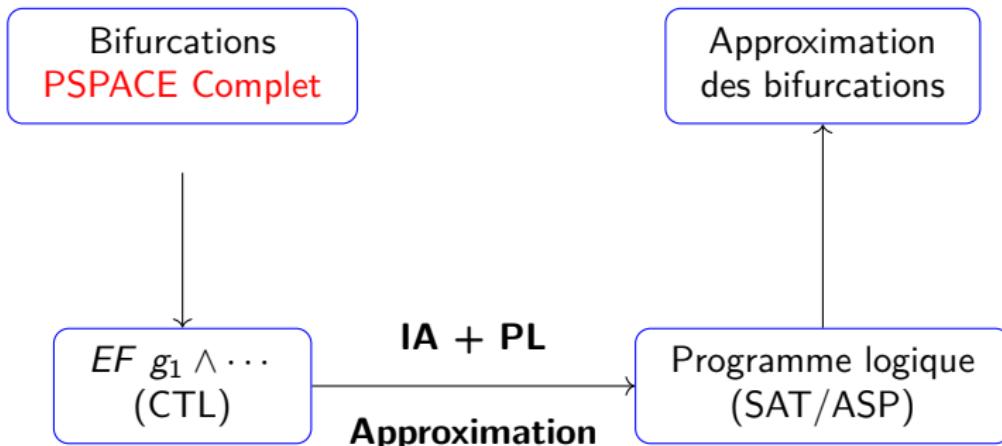


[<https://www.systembio.com/stem-cell-research/differentiation-reporters/overview>]

- Différenciation \implies Perte de capacité
- Quelles transitions (opérations) sont responsables des **bifurcations** ?
- À partir de quels états ?

- Analyse statique des bifurcations

Approximation des bifurcations

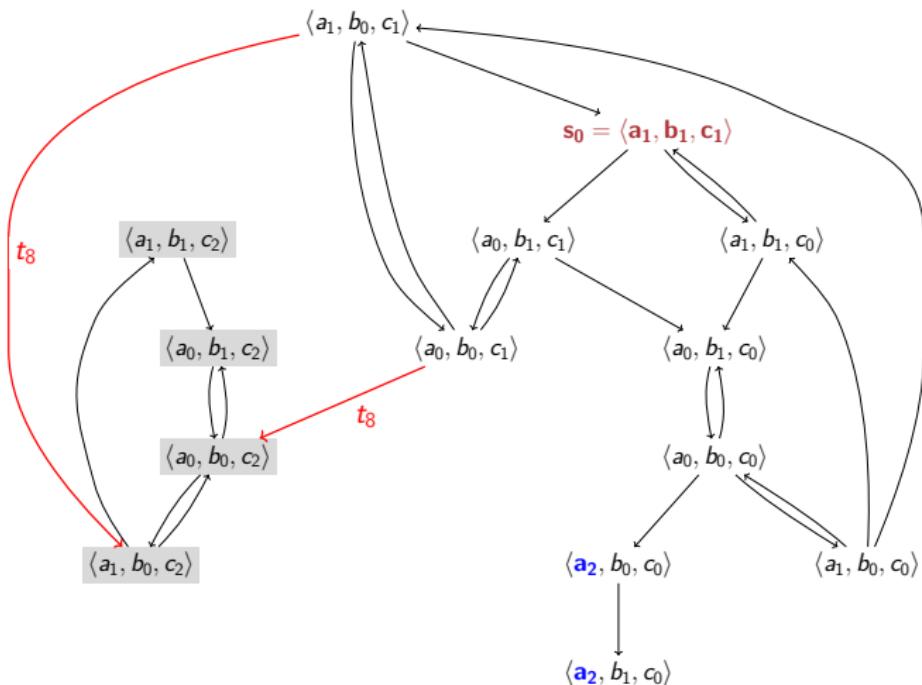


Contribution :

$\left\{ \begin{array}{l} \text{PL (Programmation logique)} \\ + \\ \text{IA (Interprétation abstraite)} \end{array} \right.$ \implies **approximations formelles des bifurcations**

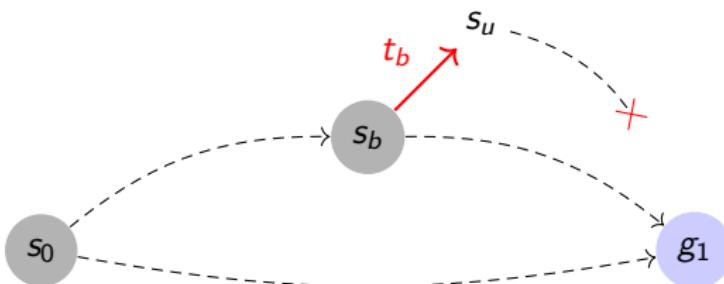
- Analyse statique des bifurcations ○ C'est quoi une bifurcation?

Illustration des bifurcations



- Analyse statique des bifurcations ○ C'est quoi une bifurcation?

Définition d'une bifurcation



Définition

t_b est une transition de bifurcation depuis s_0 pour g_1

\iff s'il existe s_b tels que:

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

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

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

Approximation formelle de l'accessibilité

Analyse statique par interprétation abstraite

Les approximations pour l'accessibilité dans les AN

[Paulevé, Magnin, Roux in *Mathematical Structures in Computer Science*, 2012]

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

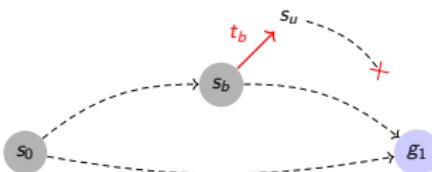
- OA (**sur-approximations**): conditions nécessaires pour $s \rightarrow^* s'$
- UA (**sous-approximation**): conditions suffisantes pour $s \rightarrow^* s'$

Intérêt:

- Éviter l'explosion de l'espace d'états
- Décider de façon efficace les propriétés d'accessibilité

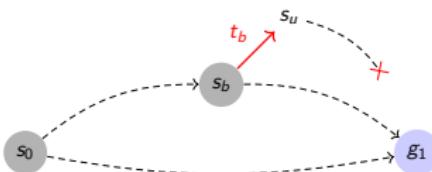
- Analyse statique des bifurcations ○ Comment identifier les bifurcations?

Relaxation du problème de l'identification des bifurcations



#: conditions suffisantes.

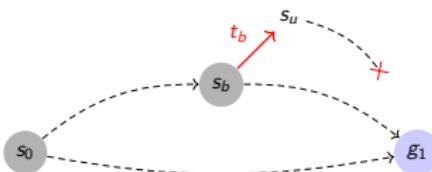
Relaxation du problème de l'identification des bifurcations



#: conditions suffisantes.

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

Relaxation du problème de l'identification des bifurcations

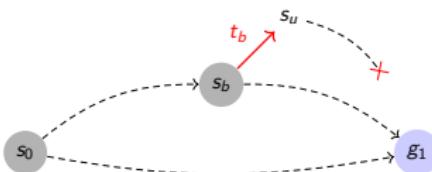


#: conditions suffisantes.

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

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

Relaxation du problème de l'identification des bifurcations



#: conditions suffisantes.

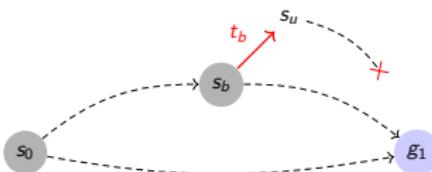
$$(C1) \quad s_b \cdot t_b \not\rightarrow^* g_1 \qquad (I1^\#) \quad \neg \text{OA}(s_u \rightarrow^* g_1) \qquad (I1^\#) \Rightarrow (C1)$$

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

$$(C3) \quad s_0 \rightarrow^* s_b \qquad (I3) \quad s_b \in \text{reach}(s_0) \qquad (I3) \Leftrightarrow (C3)$$

$$(I3^\#) \quad \text{UA}(s_0 \rightarrow^* s_b) \qquad (I3^\#) \Rightarrow (C3)$$

Relaxation du problème de l'identification des bifurcations



#: conditions suffisantes.

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

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

$$(C3) \quad s_0 \rightarrow^* s_b \qquad (I3) \quad s_b \in \text{reach}(s_0) \qquad (I3) \Leftrightarrow (C3)$$

$$(I3^\#) \quad \text{UA}(s_0 \rightarrow^* s_b) \qquad (I3^\#) \Rightarrow (C3)$$

(I1[#]) et (I2[#]) et ((I3) ou (I3[#])) $\Rightarrow t_b$ est une bifurcation.

Implémentation

Approximations formelles de l'accéssibilité

- Calcul du **Graphe de Causalité Locale**
- OA/UA: motifs particuliers dans le GLC
- Taille GLC: $\text{poly}(\#\text{automata}), \exp(|\text{single automaton}|)$

Implémentation en ASP (Answer Set Programming)

- Modélisation du problème: trouver s_b, t_b tels que
 $\neg \text{OA}(s_b \cdot t_b \rightarrow^* g_1)$ et $\text{UA}(s_b \rightarrow^* g_1)$ et $\text{UA}(s_0 \rightarrow^* s_b)$
- Énumération avec clingo

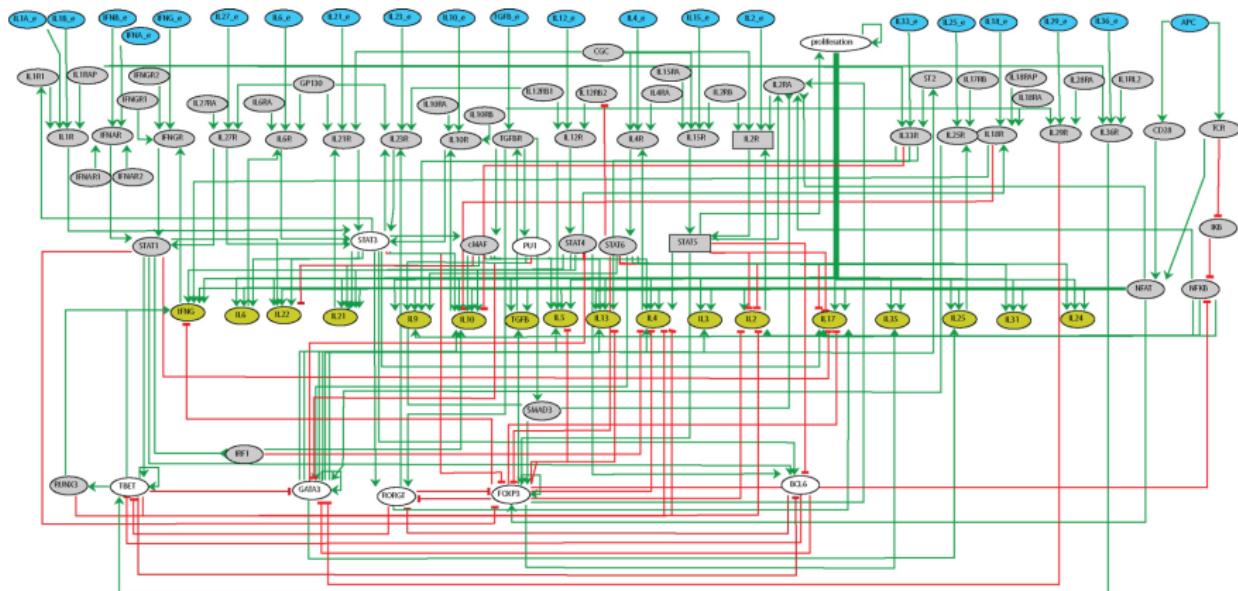
- ## o Analyse statique des bifurcations o Applications biologiques de l'identification des bifurcations

Cas d'étude

Phage Lambda : (4 composants et 11 interactions);

EGF/TNF : (28 composants et 55 interactions);

t_helper differentiation: (101 composants et 381 interactions).



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

Résultats de l'identification des bifurcations

Automata Network	Goal	M-C (NuSMV)		avec (I3)		avec (I3#)	
		$ t_b $	Time	$ t_b $	Time	$ t_b $	Time
Lambda phage $ \Sigma = 4$ $ T = 11$	Cl ₂	10	0.1s	6	0.1s	0	0.2s
	Cro ₂	3	0.1s	3	0.1s	2	0.3s
EGF/TNF $ \Sigma = 28$ $ T = 55$	NFKB ₀	5	0.2s	4	0.1s	2	0.1s
	IKB ₁	5	0.2s	3	0.1s	2	0.1s
Th_th17 $ \Sigma = 101$ $ T = 381$	RORGT ₁	18	48s	9	23s	8	26s
	BCL6 ₁	7	26s	5	23s	4	24s
Th-HTG $ \Sigma = 101$ $ T = 381$	BCL6 ₁	$\geq 2h$		$\geq 2h$		6	61s
	GATA3 ₁	$\geq 2h$		$\geq 2h$		7	34s

Implémenté en ASP (Answer Set Programming) et avec clingo 4.5.4.

Synthèse sur l'identification des bifurcations

Approximations formelles des états et transitions de bifurcation

Combinaison de deux abstractions

- $\left\{ \begin{array}{l} \text{PL (Programmation logique)} \\ + \\ \text{IA (Interprétation abstraite)} \end{array} \right.$ ⇒ approximations formelles des bifurcations
- Applicable sur les grands modèles.
- Manquer certaines bifurcations.

Application sur des modèles biologiques réels

- Lambda phage, EGF/TNF, Différenciation des lymphocytes T.
- Base pour la reprogrammation cellulaire.

Publication

- Louis Fippo Fitime, Olivier Roux, Carito Guziolowski, Loïc Paulevé. [Identification of Bifurcations in Biological Regulatory Networks using Answer-Set Programming](#). In *12th International Workshop on Constraint-Based Methods for Bioinformatics (WCB'12)*. September 2016.

Conclusions

Conclusions

Modélisation hybride des réseaux RSTC

- Formalisation des connaissances biologiques: Translation des motifs en AN
- Inférence et intégration des paramètres temporels et stochastiques
- Raffinement qualitatif et quantitatif de la dynamique

Conclusions

Modélisation hybride des réseaux RSTC

- Formalisation des connaissances biologiques: Translation des motifs en AN
- Inférence et intégration des paramètres temporels et stochastiques
- Raffinement qualitatif et quantitatif de la dynamique

Analyse statique des propriétés quantitatives (probabilité/délai)

- $\left\{ \begin{array}{l} \text{QS (Classes d'Équivalences d'États)} \\ + \\ \text{IA (Interprétation Abstraite)} \end{array} \right.$ \implies approximations formelles des propriétés quantitatives
- Estimation efficace d'une borne inférieure de la probabilité et du délai de l'accessibilité
- Faible complexité théorique

Conclusions

Modélisation hybride des réseaux RSTC

- Formalisation des connaissances biologiques: Translation des motifs en AN
- Inférence et intégration des paramètres temporels et stochastiques
- Raffinement qualitatif et quantitatif de la dynamique

Analyse statique des propriétés quantitatives (probabilité/délai)

- $\left\{ \begin{array}{l} \text{QS (Classes d'Équivalences d'États)} \\ + \\ \text{IA (Interprétation Abstraite)} \end{array} \right.$ \implies approximations formelles des propriétés quantitatives
- Estimation efficace d'une borne inférieure de la probabilité et du délai de l'accessibilité
- Faible complexité théorique

Identification des Bifurcations

- $\left\{ \begin{array}{l} \text{PL (Programmation logique)} \\ + \\ \text{IA (Interprétation abstraite)} \end{array} \right.$ \implies approximation formelle des bifurcations
- Applicable aux grands modèles (vs model-checking)

Perspectives

Perspectives

Simulation stochastique & de l'analyse statique des traces

- Analyse statistique qui prenne en compte le temps.
- Utilisation des règles d'associations pour identifier les relations non encore connues.

Perspectives

Simulation stochastique & de l'analyse statique des traces

- Analyse statistique qui prenne en compte le temps.
- Utilisation des règles d'associations pour identifier les relations non encore connues.

Analyse statique des propriétés quantitatives

- Implémentation (en cours...) d'un analyseur statique des propriétés (CCSL Logic)
- Analyse statique des comportements à long terme (états stables)
- Analyse statique des modèles paramétriques

Perspectives

Simulation stochastique & de l'analyse statique des traces

- Analyse statistique qui prenne en compte le temps.
- Utilisation des règles d'associations pour identifier les relations non encore connues.

Analyse statique des propriétés quantitatives

- Implémentation (en cours...) d'un analyseur statique des propriétés (CCSL Logic)
- Analyse statique des comportements à long terme (états stables)
- Analyse statique des modèles paramétriques

Identification des Bifurcations

- Sur-approximation des bifurcations
- Utilisation des bifurcations pour l'analyse des probabilités
- Applications pour **prédire les cibles pour la reprogrammation cellulaire.**

Soutenance de thèse de doctorat

Modélisation hybride, Analyse et Vérification des réseaux de régulation biologique

Louis Fippo Fitime

louis.fippo-fitime@irc-cyn.ec-nantes.fr

Lundi 28 novembre 2016

Rapporteurs: Christine FROIDEVAUX, Professeur des universités, Université Paris Sud
Simon de GIVRY, Chargé de recherche INRA, INRA Toulouse

Examinateurs: Anne POUPON, Directrice de recherche CNRS, INRA de Tours
Sylvain SOLIMAN, Chargé de recherche INRIA, INRIA Paris Saclay

Directeur de thèse: Olivier ROUX, Professeur des universités, École centrale de Nantes

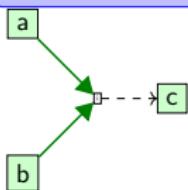
Co-encadrant de thèse: Carito GUZIOLOWSKI, Maître de conférences, École centrale de Nantes

Merci de votre attention!

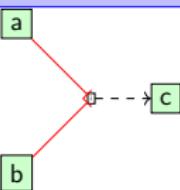
Sortes de synchronisation

Motifs biologiques

activations indep.

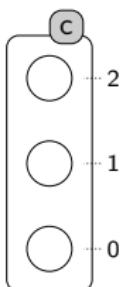
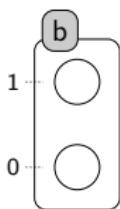
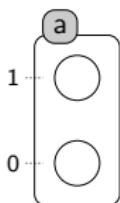


inhibitions indep.



Sortes de synchronisation

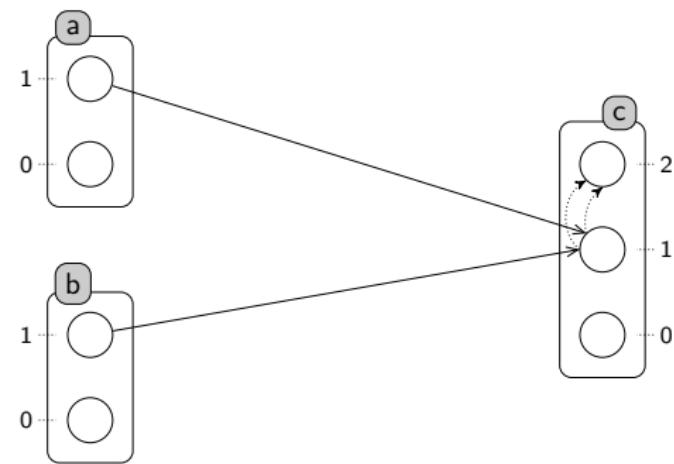
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

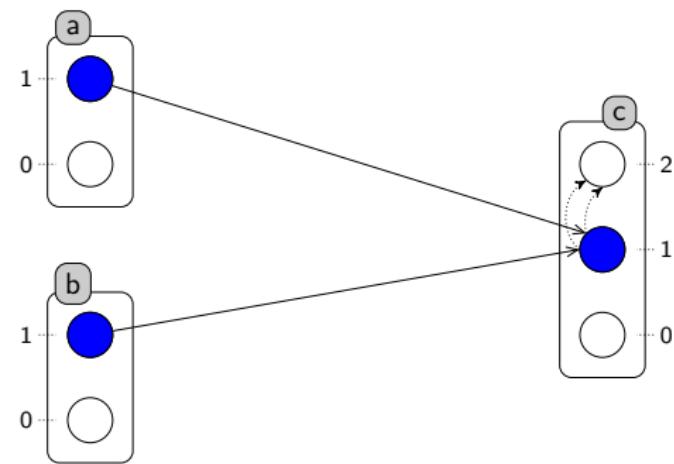
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

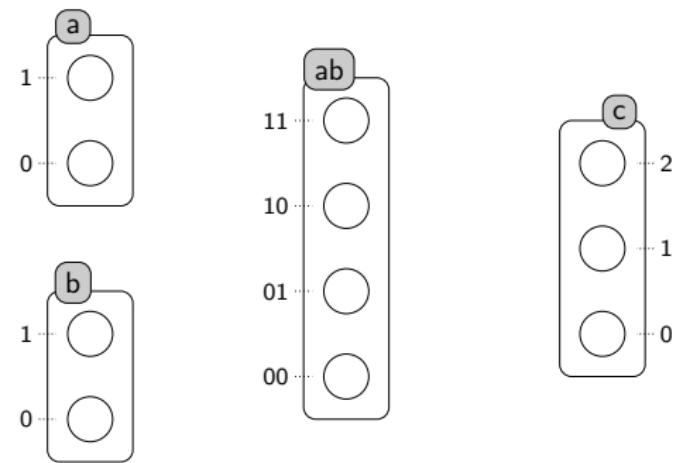
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

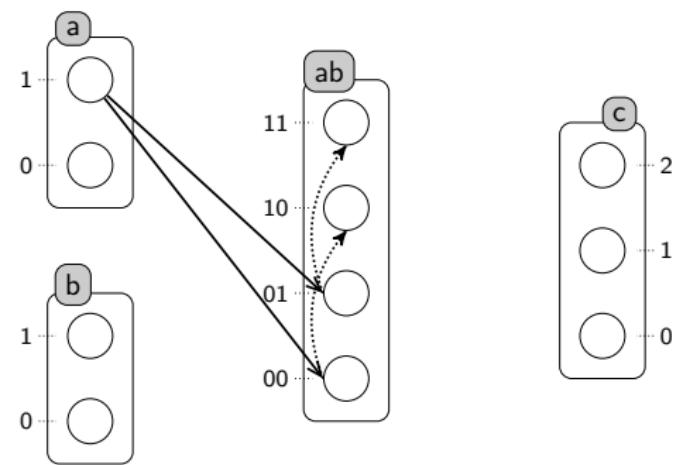
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

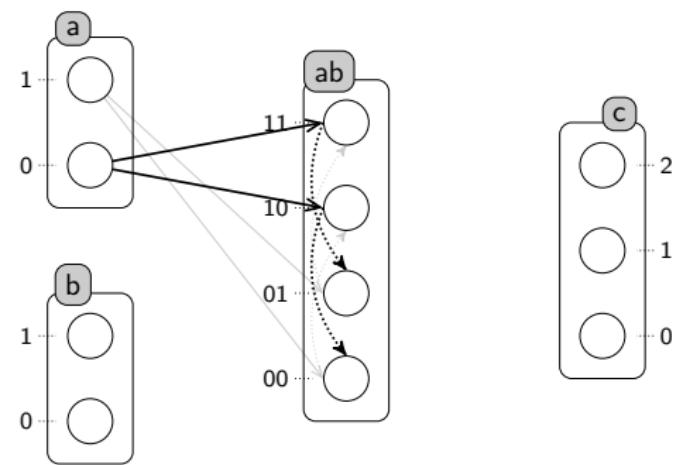
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

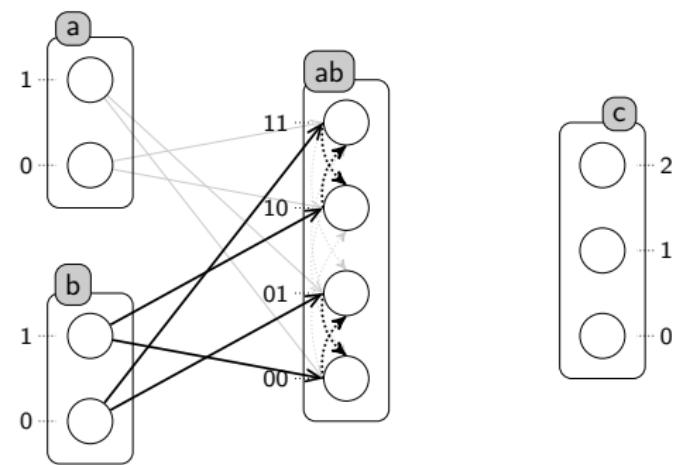
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

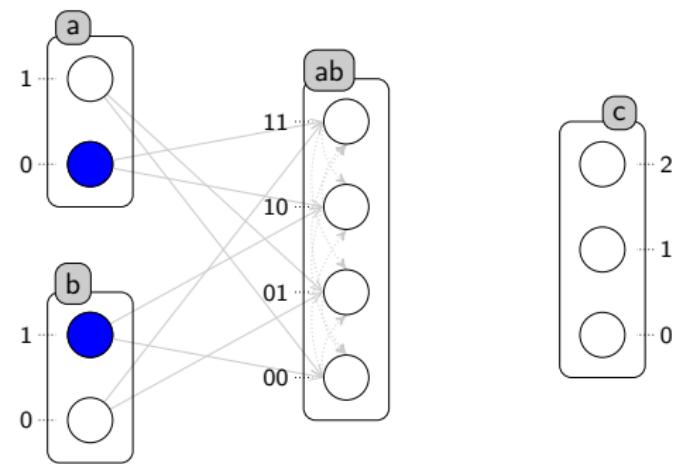
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

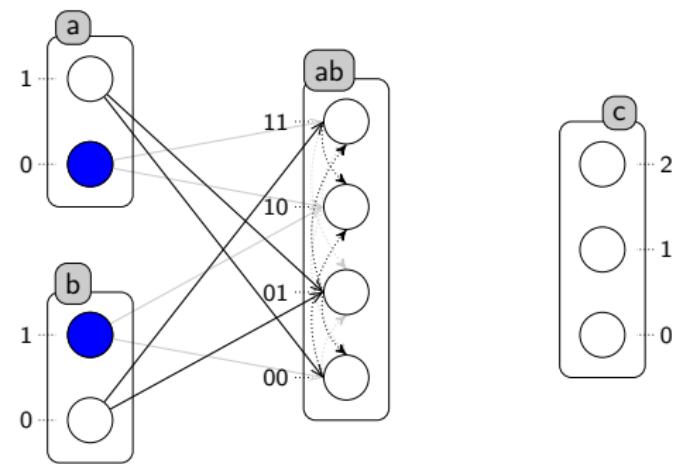
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

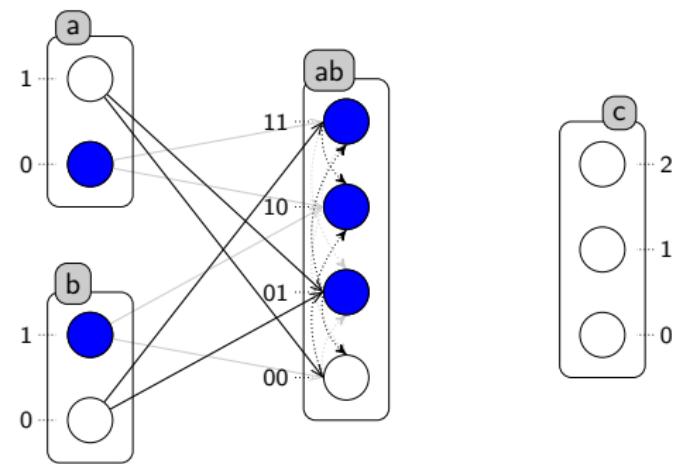
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

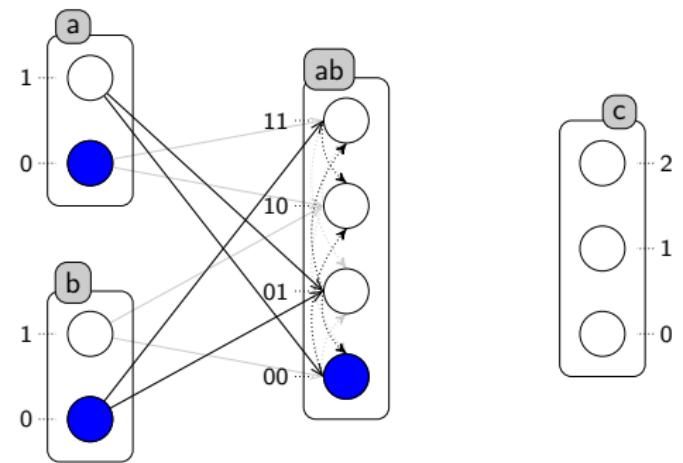
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

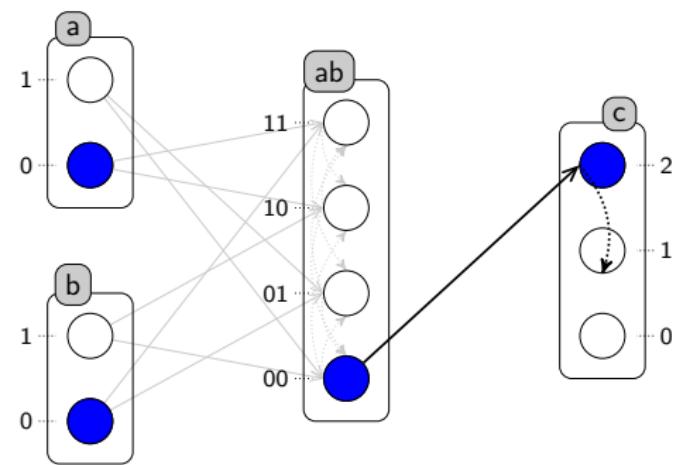
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

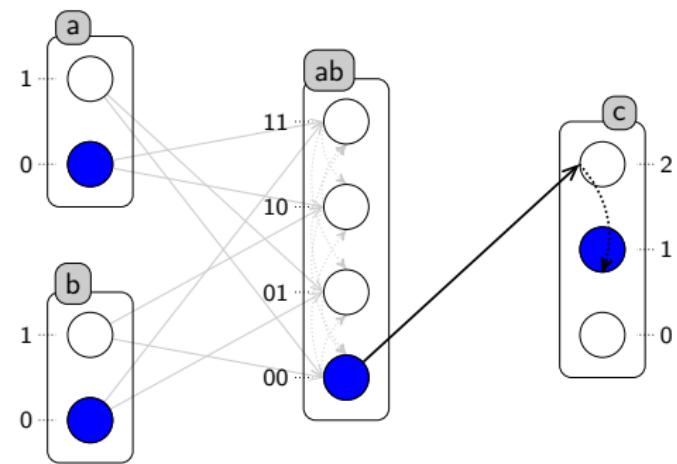
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

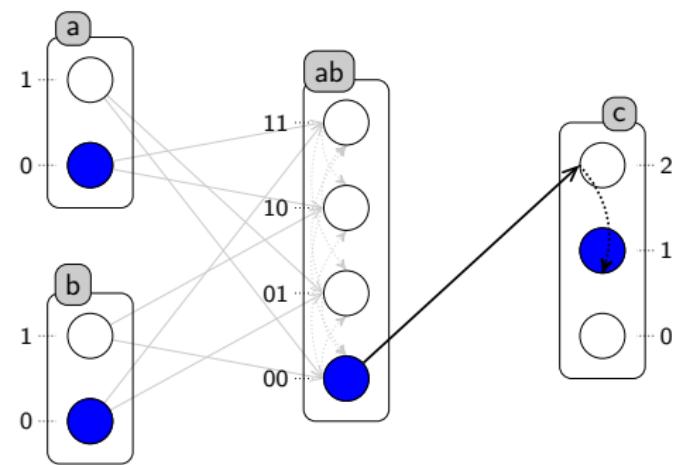
Raffinement structurel de la dynamique



Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

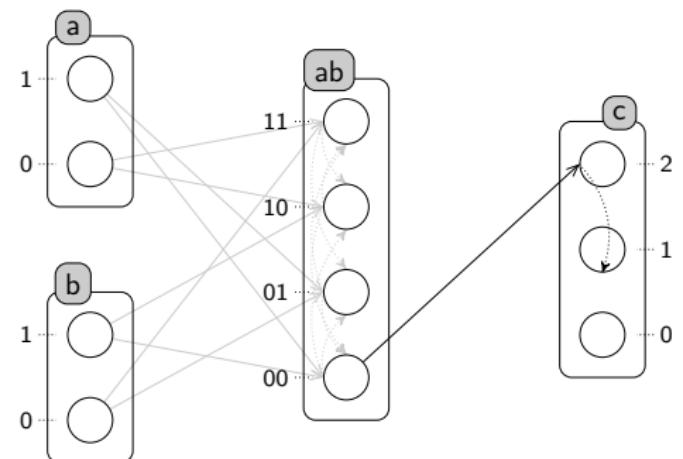
Raffinement structurel de la dynamique



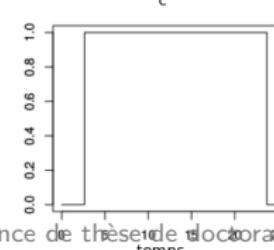
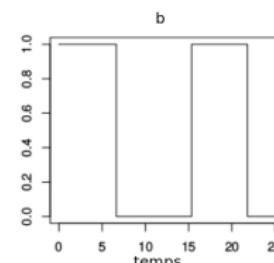
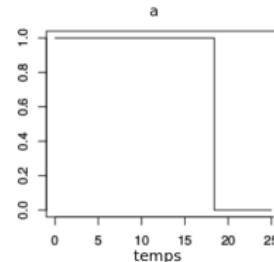
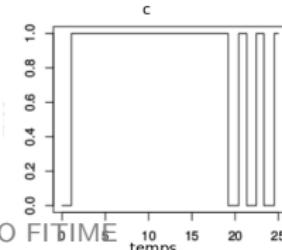
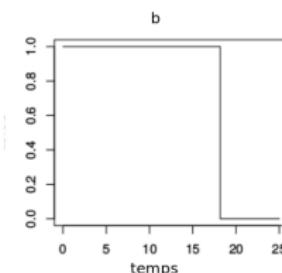
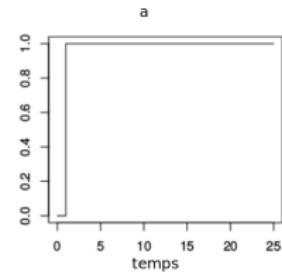
Comment introduire la **Synchronisation** entre les sortes? $a_0 \wedge b_0 \rightarrow c_2 \uparrow c_1$

Sortes de synchronisation

Raffinement structurel de la dynamique



Résultat avec la synchronisation



Graphe de causalité locale quantifié

- Probabilité d'un **nœud état local**:

$$\mathcal{P}(\eta_k, \varsigma, [\mathcal{G}_\varsigma^\omega]) \stackrel{\Delta}{=} \sum_{\eta_{k-1,i} \in Ob_{j_k}} 1_\varsigma(\text{cible}(\eta_{k-1})) \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, [\mathcal{G}_\varsigma^\omega])$$

Graphe de causalité locale quantifié

- Probabilité d'un **nœud état local**:

$$\mathcal{P}(\eta_k, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil) \stackrel{\Delta}{=} \sum_{\eta_{k-1,i} \in Obj_k} 1_\varsigma(\text{cible}(\eta_{k-1})) \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil)$$

- Probabilité d'un **nœud objectif**:

$$\mathcal{P}(\eta_k, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil) \stackrel{\Delta}{=} \frac{\sum_{\eta_{k-1,i} \in Sol_k} r_{k-1} \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil)}{\sum_{\eta_{k-1,i} \in Sol_k} r_{\eta_{k-1}} \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil) + \sum_{j \in \bar{Sol}_k} r_j} \cdot \mathcal{P}_{(A_{\text{cible}(\eta_k)}, *)}(t)$$

Graphe de causalité locale quantifié

- Probabilité d'un **nœud état local**:

$$\mathcal{P}(\eta_k, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil) \stackrel{\Delta}{=} \sum_{\eta_{k-1,i} \in Obj_k} 1_\varsigma(\text{cible}(\eta_{k-1})) \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil)$$

- Probabilité d'un **nœud objectif**:

$$\mathcal{P}(\eta_k, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil) \stackrel{\Delta}{=} \frac{\sum_{\eta_{k-1,i} \in Sol_k} r_{k-1} \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil)}{\sum_{\eta_{k-1,i} \in Sol_k} r_{k-1} \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil) + \sum_{j \in \bar{Sol}_k} r_j} \cdot \mathcal{P}_{(A_{\text{cible}(\eta_k)}, *)}(t)$$

- Probabilité d'un **nœud Solution**:

$$\mathcal{P}(\eta_k, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil) \stackrel{\Delta}{=} \prod_{\eta_{k-1,i} \in LS_k} \mathcal{P}(\eta_{k-1,i}, \varsigma, \lceil \mathcal{G}_\varsigma^\omega \rceil)$$

Graphe de causalité locale quantifié

- Probabilité d'un **nœud état local**:

$$\mathcal{P}(\eta_k, \varsigma, [\mathcal{G}_\varsigma^\omega]) \stackrel{\Delta}{=} \sum_{\eta_{k-1,i} \in Obj_k} 1_\varsigma(\text{cible}(\eta_{k-1})) \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, [\mathcal{G}_\varsigma^\omega])$$

- Probabilité d'un **nœud objectif**:

$$\mathcal{P}(\eta_k, \varsigma, [\mathcal{G}_\varsigma^\omega]) \stackrel{\Delta}{=} \frac{\sum_{\eta_{k-1,i} \in Sol_k} r_{k-1} \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, [\mathcal{G}_\varsigma^\omega])}{\sum_{\eta_{k-1,i} \in Sol_k} r_{\eta_{k-1}} \cdot \mathcal{P}(\eta_{k-1,i}, \varsigma, [\mathcal{G}_\varsigma^\omega]) + \sum_{j \in \bar{Sol}_k} r_j} \cdot \mathcal{P}_{(A_{\text{cible}(\eta_k)}, *)}(t)$$

- Probabilité d'un **nœud Solution**:

$$\mathcal{P}(\eta_k, \varsigma, [\mathcal{G}_\varsigma^\omega]) \stackrel{\Delta}{=} \prod_{\eta_{k-1,i} \in LS_k} \mathcal{P}(\eta_{k-1,i}, \varsigma, [\mathcal{G}_\varsigma^\omega])$$

