LIPN seminar

Hybrid Modelling, Analysis and Verification of large-scale Biological Networks

Louis Fippo Fitime\textsuperscript{1,3}

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

Joint work with:

Olivier Roux\textsuperscript{1}, Carito Guziolowski\textsuperscript{1} and Loïc Paulevé\textsuperscript{2}

\textsuperscript{1} MeForBio / IRCCyN / École Centrale de Nantes (Nantes, France)
\textsuperscript{2} Bioinfo / LRI / CNRS (Paris, France)
\textsuperscript{3} LCR / LIPN / CNRS (Paris, France)

olivier.roux@irccyn.ec-nantes.fr

carito.guziolowski@irccyn.ec-nantes.fr

loic.pauleve@lri.fr
Cellular processes → networks of biological interactions.

Nodes (biological components), edges (interactions).
Contributions

Dynamics abstraction level

Properties precision

Generals properties: bounds; nbr of attractors; functionality, ...

Stable states

Reachability

Bifurcations

(CTL)

Quantitative

reachability

(CCSL)

Quantitative dynamics analysis

Static analysis

Dynamics analysis

Chronometry

Chronology
Introduction

Contributions

- Generals properties: bounds
  - nbr of attractors;
  - functionality,...

- Stable states
- Reachability
  - Bifurcations
    - (CTL)

- Quantitative reachability
  - (CCSL)

Dynamics abstraction level

Properties precision
Contributions

Generals properties: bounds
  nbr of attractors;
  functionality,...

Stable states

Reachability
Bifurcations
(CTL)

Quantitative
reachability
(CCSL)
Contributions

- Generals properties: bounds; nbr of attractors; functionality, ...
- Stable states
- Reachability (CTL)
- Bifurcations
- Quantitative reachability (CCSL)

Dynamics abstraction level

- Static analysis
- Dynamics analysis
- Static analysis
- Static analysis

Properties precision

- Chronology
- Chronometry
- Quantitative dynamics analysis
Introduction

Contributions

Dynamics abstraction level

General properties: bounds
nbr of attractors; functionality,...

Stable states

Reachability
Bifurcations
(CTL)

Quantitative
reachability
(CCSL)

Chronology

Chronometry

Properties precision
Contributions

General properties: bounds
nbr of attractors; fonctionality, ...

Stable states

Bifurcations
Reachability (CTL)

Qualitative
Quantitative

Quantitative reachability (CCSL)
Introduction

Contributions

Dynamics abstraction level

Properties precision

General properties: bounds
  nbr of attractors;
  fonctionality,...

Stable states

Bifurcations

Reachability

(CTL)

Quantitative reachability

(CCSL)

Chronometry

Quantitative

Qualitative

Automata Networks

Static analysis

1. A

Modelling

PINT

Chronology

1. B

Static analysis

Simulation

1. A

Analysis

Static analysis

Modelling

Louis FIPPO FITIME 5/48 LIPN seminar — 09/03/2017
Introduction

Contributions

Dynamics abstraction level

Properties precision

General properties: bounds, number of attractors; functionality, ...

Stable states

Bifurcations

Reachability (CTL)

Quantitative reachability (CCSL)

Automata Networks

Static analysis

Chronology

Chromotometry

Modelling

PINT

Analysis

Simulation

1.A

1.B

Qualitative

Quantitative

LIPN seminar — 09/03/2017
Introduction

Contributions

- Dynamics abstraction level
- Properties precision
  - General properties: bounds, number of attractors, functionality, ...
  - Stable states
  - Bifurcations
  - Reachability (CTL)
  - Quantitative reachability (CCSL)

Automata Networks

- Static analysis
- Chronometry
  - Chronology
- PINT

1. Analysis
  - 1.A: Modelling
  - 1.B: Simulation

2. Static analysis

Qualitative

Quantitative

Louis FIPPO FITIME 5/48 LIPN seminar — 09/03/2017
Contributions

General properties: bounds
nbr of attractors; fonctionality,...

Stable states

3

Bifurcations
Reachability
(CTL)

1.A
Simulation
Modeling
PINT

1.B
Analysis
Simulation

Static analysis

2

Static analysis

Qualitative
Quantitative

Quantitative
reachability
(CCSSL)

Static analysis

Chronology
Chronometry

Dynamics abstraction level

Properties precision
Introduction

Contributions

Dynamics abstraction level

Properties precision

General properties: bounds
- nbr of attractors;
- functionality,...

Stable states

Bifurcations

Reachability (CTL)

Quantitative reachability (CCSL)

Automata network

Chronsometry

Static analysis

Modeling

PINT

Qualitative

Quantitative

Simulation

Analysis

1.A

1.B

2

3

6/48 LIPN seminar — 09/03/2017
Biological networks modelling

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

$\langle a_0, b_1, c_0 \rangle$

$\langle a_2, b_0, c_0 \rangle$

$\langle a_1, b_0, c_0 \rangle$

$\langle a_2, b_1, c_0 \rangle$
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

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

$s_0 = \langle a_0, b_0, c_0 \rangle$
**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 $t_1 = a_0 \xrightarrow{b_0} a_1$, $t_2 = a_1 \rightarrow a_0$, $t_3 = a_0 \xrightarrow{b_0, c_0} a_2$, $t_4 = b_0 \rightarrow b_1$

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

$\langle a_1, b_0, c_0 \rangle$
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
\[
\begin{align*}
t_1 &= a_0 \xrightarrow{b_0} a_1, \\
t_2 &= a_1 \rightarrow a_0, \\
t_3 &= a_0 \xrightarrow{b_0, c_0} a_2, \\
t_4 &= b_0 \rightarrow b_1
\end{align*}
\]
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: components \( a, b, c \)

local states: levels of expression \( c_0, c_1, c_2 \)

States: sets of active local states

Transitions: dynamics

\[ t_1 = a_0 \xrightarrow{b_0} a_1, \quad t_2 = a_1 \to a_0, \quad t_3 = a_0 \xrightarrow{b_0,c_0} a_2, \quad t_4 = b_0 \to b_1 \]
Stochastic Features

[Paulevé et al. 2010]

- Introduces time features
- Parameters: either $(r, sa)$, or the firing interval $[d; D]$.
  - Tests by simulation

![Diagram of stochastic features](image)

action duration
**Introduction**

**Automata Network**

- **Contributions**
  - **Dynamics abstraction level**
  - **Properties precision**
  - **General properties**: bounds, nbr of attractors; functionality, ...

**Automata network**

- **Static analysis**
  - Stable states
  - Bifurcations
  - Reachability (CTL)
  - Quantitative reachability (CCSL)

**Chronometry**

- **Qualitative**
  - Simulation
  - Analysis

**Modelling**

- **Simulation**
  - PINT

**Louis FIPPO FITIME 10/48 LIPN seminar — 09/03/2017**
Introduction

Automata Network

Contributions

General properties: bounds

- nbr of attractors;
- functionality,...

Stable states

Bifurcations

Reachability

(CTL)

Quantitative reachability

(CCSL)

Chronometry

Quantitative

Qualitative

PINT

Static analysis

Simulation

Analysis

Modelling

1.A

1.B

2

3

Chronology

Properties precision

Louis FIPPO FITIME 10/48 LIPN seminar — 09/03/2017
Dynamics refinement by Time Series Data Integrating

<table>
<thead>
<tr>
<th>General knowledge</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature</td>
<td></td>
</tr>
<tr>
<td>Hypotheses</td>
<td></td>
</tr>
</tbody>
</table>

Times series data

<table>
<thead>
<tr>
<th>Genes</th>
<th>1h</th>
<th>...</th>
<th>24h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 1</td>
<td></td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>Gene 2</td>
<td></td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

```
a

p1?

b

p2?

c
```
Dynamics refinement by Time Series Data Integrating

<table>
<thead>
<tr>
<th>General knowledge</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature</td>
<td></td>
</tr>
<tr>
<td>Hypotheses</td>
<td></td>
</tr>
</tbody>
</table>

**Motivation**

**Algebraic Modelling**

- Formal characterisation of topology and dynamics of biological regulatory networks.
- Time-series data integration.
- Qualitative & Quantitative refinement of the dynamics.
- Stochastic simulation and statistic analysis.

**Times series data**

<table>
<thead>
<tr>
<th>Genes</th>
<th>1h</th>
<th>...</th>
<th>24h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene 1</td>
<td></td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>Gene 2</td>
<td></td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>
RSTC Network

multi-layer receptor-signaling-transcription-cell state

- Pathway Interaction Database
- **293 nodes**: signaling proteins, transcription factors, mRNA expressions
- **375 interactions**: activations, inhibitions, complexes dissociation
Dynamics refinement by Time Series Data Integrating Biological Networks

RSTC Network
multi-layer receptor-signaling-transcription-cell state
Dynamics refinement by Time Series Data Integrating Biological Networks

RSTC Network
multi-layer receptor-signaling-transcription-cell state
Biological Network formalization

Automatic patterns detection and formalization
Biological Network formalization

Automatic patterns detection and formalization
Biological Network formalization

Automatic patterns detection and formalization

Biological network

RSTC network

Patterns Detection

in RSTC network

Set of patterns

of RSTC network

Biological Patterns

Simple activation

Simple inhibition

activation or inhibition

Louis FIPPO FITIME
Dynamics refinement by Time Series Data Integrating Biological Networks

**Biological Network formalization**

Automatic patterns detection and formalization

---

**Biological Patterns**

- **Simple activation**
  
  ![Diagram of simple activation](image)

- **Simple inhibition**
  
  ![Diagram of simple inhibition](image)

- **Activation or inhibition**
  
  ![Diagram of activation or inhibition](image)

---

**Time delay**

Estimate from TSD

---

**Patterns Detection in RSTC network**

**Set of patterns of RSTC network**

**Patterns translation in AN (PH) model**
**Biological Network formalization**

Automatic patterns detection and formalization

- **Biological network**
  - RSTC network

- **Patterns Detection**
  - in RSTC network

- **Set of patterns**
  - of RSTC network

- **Patterns translation**
  - in AN (PH) model

- **AN (PH) Patterns**

**Biological Patterns**

- **Simple activation**
  - a → b

- **Simple inhibition**
  - a → b

- **Activation or inhibition**
  - a → c → b

**AN (Process Hitting) Transformations**

- Time delay
  - Estimate from TSD

- Set of patterns
  - of RSTC network

- Patterns translation
  - in AN (PH) model

- AN (PH) Patterns
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Dynamics refinement by Time Series Data Integrating Biological Networks

Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states,...

Transient Nodes: translocations, modifications, transcriptions,...

Ecad

PS6 cplx4 PS8 cplx6 PS9

PS1 cplx3 PS2

cplx5

PS3 cplx1 PS4

SN1

a

b

c

d

SN2

PS2

PS3

PS3

a

b

c
Biological Network formalization

Automatic patterns detection (method)

Terminal Nodes: proteins, complexes, mRNA expressions, cellular states, ...

Transient Nodes: translocations, modifications, transcriptions, ...

Complexity: The algorithm has a time complexity of $O(|V| \log(h))$. Where $|V|$ is the set of nodes and $|E|$ is the set of edges.

$h$ is the average height of the patterns in the RSTC network. In the worst case $h = \log(|V|)$. 
Dynamics refinement by Time Series Data Integrating Stochastic Parameters Estimation & Integration

Time series data

- Experiment: calcium stimuli
- Measured at 10 time-points (0-24hrs)
- 200 transcripts selected (dynamic patterns)
- We included in our model a subset of 12 of them
Dynamics refinement by Time Series Data Integrating ○ Stochastic Parameters Estimation & Integration

Data integration
Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to its TSD is:

\[ r_i = \frac{1}{t_i - t_{i-1}} \]
Dynamics refinement by Time Series Data Integrating

Stochastic Parameters Estimation & Integration

Data integration

Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to its TSD is:

\[ r_i = \frac{1}{t_i - t_{i-1}} \]

Louis FIPPO FITIME 18/48 LIPN seminar — 09/03/2017
Dynamics refinement by Time Series Data Integrating

Stochastic Parameters Estimation & Integration

Data integration

Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to it TSD is

\[ r_i = \frac{1}{t_i - t_{i-1}} \]
Dynamics refinement by Time Series Data Integrating ○ Stochastic Parameters Estimation & Integration

Data integration
Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to its TSD is:

\[ r_i = \frac{1}{t_i - t_{i-1}} \]

where:

- \( t_i \) is the time at which the rate is estimated.
- \( t_{i-1} \) is the previous time.

The graph shows a time series with points marked at different times, indicating the maxima (Max) and minima (Min) of the data. The parameters th1 and th2 are used to analyze the data points.

The diagram includes a vertical line at each time point, indicating the transitions between states.
Dynamics refinement by Time Series Data Integrating o Stochastic Parameters Estimation & Integration

Data integration
Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to its TSD is

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

$$a_1 \rightarrow b_0 \uparrow b_1 \text{ with } r_1 = \frac{1}{t_1 - t_0}$$
Dynamics refinement by Time Series Data Integrating
Stochastic Parameters Estimation & Integration

Data integration
Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to its TSD is

\[ r_i = \frac{1}{t_i - t_{i-1}} \]

\[ a_1 \rightarrow b_0 \uparrow b_1 \text{ with } r_1 = \frac{1}{t_1 - t_0} \]

\[ a_1 \rightarrow b_1 \uparrow b_2 \text{ with } r_2 = \frac{1}{t_2 - t_1} \]
Dynamics refinement by Time Series Data Integrating

Data integration

Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to its TSD is:

\[ \frac{1}{t_i - t_{i-1}} \]

Examples:

- \( a_1 \rightarrow b_0 \uparrow b_1 \) with \( r_1 = \frac{1}{t_1 - t_0} \)
- \( a_1 \rightarrow b_1 \uparrow b_2 \) with \( r_2 = \frac{1}{t_2 - t_1} \)
- \( a_0 \rightarrow b_2 \uparrow b_1 \) with \( r_3 = \frac{1}{t_3 - t_2} \)
Dynamics refinement by Time Series Data Integrating ○ Stochastic Parameters Estimation & Integration

Data integration
Parameters inference/Principe

The formula to estimate the rate of the dynamics of a component according to its TSD is $r_i = \frac{1}{t_i - t_{i-1}}$

$a_1 \rightarrow b_0 \nrightarrow b_1$ with $r_1 = \frac{1}{t_1 - t_0}$

$a_1 \rightarrow b_1 \nrightarrow b_2$ with $r_2 = \frac{1}{t_2 - t_1}$

$a_0 \rightarrow b_2 \nrightarrow b_1$ with $r_3 = \frac{1}{t_3 - t_2}$

$a_0 \rightarrow b_1 \nrightarrow b_0$ with $r_4 = \frac{1}{t_4 - t_3}$
Contributions

General properties: bounds
- nbr of attractors;
- fonctionality,...

Stable states

Bifurcations
Reachability
(CTL)

Quantitative
reachability
(CCRL)

Chronometry
Chronology

Properties precision

Modelling
Analysis
Simulation

Static analysis

1A

PINT

1.B

2

3
Dynamics refinement by Time Series Data Integrating - Biological Application: Simulations & Validation

Contributions

General properties: bounds
- nbr of attractors;
- funcionality,...

Stable states

Bifurcations
Reachability
(CTL)

PINT

Qualitative Quantitative
Reachability
(CCCL)

Chronometry

Modeling

1.A Simulation

Static analysis

3

Analyse

1.B Simulation

Static analysis

2

Louis FIPPO FITIME 19/48 LIPN seminar — 09/03/2017
Hypothesis of simulation

Structural and dynamics Hypothesis

- **E_cadherin**: has a pulse signal
- We introduce **auto-hit** at the transcription factor level
- We assume that mRNA expression have **three levels of expression**

Diagram:
- Ecad
- PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9
- cplx1, cplx2, cplx3, cplx4, cplx5, cplx6
- SN1, SN2
- a, b, c, d
Simulations and analysis

- Dynamics refinement by Time Series Data Integrating
- Biological Application: Simulations & Validation

For $r_a = r_i = 10.0$ and $s_a = 50$ for all signalling proteins with estimated $r$ and $s_a$ for MKP3, MKP1, uPAR, Hes5, ... according to their expression profiles.
Simulations and analysis

- Input node of the system (E_cadherin)
- For biological processes (Cell adhesion, Cell cycle arrest, Keratinocyte differentiation)
Simulation and Trace analysis

for each component $C_i$, $1 \leq i \leq P$, $N$ simulations will generate $\omega_{i1}, \omega_{i2}, \ldots, \omega_{iN}$ words.
for $1 \leq j \leq N$

$$\omega_{ij} \Rightarrow A_{C_i} \Rightarrow \text{yes/no}$$

$$\% \text{ of Acceptance} = \frac{|YES|}{|Simulations|}$$
### Simulation and Trace analysis

<table>
<thead>
<tr>
<th>Automate</th>
<th>components</th>
<th>% validation</th>
<th>% of acceptance $T_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{A}_2(01210)$</td>
<td>A20</td>
<td>91</td>
<td>100</td>
</tr>
<tr>
<td>$\mathcal{A}_2(01210)$</td>
<td>IL1_beta</td>
<td>81</td>
<td>100</td>
</tr>
<tr>
<td>$\mathcal{A}_2(01210)$</td>
<td>IL8</td>
<td>93</td>
<td>100</td>
</tr>
<tr>
<td>$\mathcal{A}_2(01210)$</td>
<td>TNF_alpha</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\mathcal{A}_3(01211)$</td>
<td>uPar</td>
<td>76</td>
<td>99</td>
</tr>
<tr>
<td>$\mathcal{A}_3(01211)$</td>
<td>ET1</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>$\mathcal{A}_4(0121210)$</td>
<td>DKK1</td>
<td>13</td>
<td>43</td>
</tr>
<tr>
<td>$\mathcal{A}_5(0121211)$</td>
<td>Hes5</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>$\mathcal{A}_5(0121211)$</td>
<td>MKP1</td>
<td>9</td>
<td>97</td>
</tr>
<tr>
<td>$\mathcal{A}_6(0212)$</td>
<td>SM22</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>$\mathcal{A}_7(02010)$</td>
<td>MKP3</td>
<td>11</td>
<td>98</td>
</tr>
<tr>
<td>$\mathcal{A}_8(02121)$</td>
<td>Tfr</td>
<td>0</td>
<td>94</td>
</tr>
</tbody>
</table>
Static analysis of quantitative properties in Stochastic Automata Networks

Contributions

- Dynamics abstraction level
- Properties precision

**General properties:** bounds
  nbr of attractors; fonctionality,...

- Stable states
- Bifurcations
- Reachability (CTL)
- Quantitative reachability (CCSL)

**Automata network**

Modeling
1.A

Simulation
1.B

Analysis

PINT

Chronometry
Chronology

Qualitative
Quantitative
Static analysis of quantitative properties in Stochastic Automata Networks

Contributions

- Dynamics abstraction level
- Properties precision

Automata network

General properties: bounds
  - nbr of attractors;
  - functionality,

Stable states

Bifurcations
Reachability (CTL)

Quantitative
Reachability (CCSL)

Chronometry

PINT

1.A
Modelling
Simulation

1.B
Analysis

2
Static analysis

3
Static analysis

3
Static analysis

Qualitative

Quantitative
Static analysis of quantitative properties in Stochastic Automata Networks

Stochastic 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
$t_1 = a_0 \xrightarrow{b_0} a_1$,  $t_2 = a_1 \rightarrow a_0$,  $t_3 = a_0 \xrightarrow{b_0, c_0} a_2$,  $t_4 = b_0 \rightarrow b_1$
Stochastic 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 $t_1 = a_0 \xrightarrow{b_0,2} a_1$, $t_2 = a_1 \rightarrow a_0$, $t_3 = a_0 \xrightarrow{b_0,c_0,2} a_2$, $t_4 = b_0 \rightarrow b_1$
Static analysis of quantitative properties in Stochastic Automata Networks

**SAN as CTMC**

Bad news: state space explosion
SAN as CTMC

Probability to have a transition from state $s$:

$$P_{(s, \ast)}(t) = 1 - e^{-\lambda t}$$

Probability to have a transition from $s$ to $s'$:

$$P_{(s, s')} (t) = \frac{R(s, s')}{E(s)} \cdot P_{(s, \ast)}(t)$$

Probability to have a transition from $s$ to a set of states:

$$P_{(s, A)} (t) = \frac{R(s, A)}{E(s)} \cdot P_{(s, \ast)}(t)$$

Bad news: state space explosion
Aim:

- Formal characterization of quantitative properties.
- Static analysis of Probability and Delay of reachability properties.
- Estimate a lower bound and an upper bound of a quantitative properties.

Contribution:

\[ \begin{align*}
0 \quad \text{binf}_{\mathcal{P}(\mathcal{R})} \quad \mathcal{P}(\mathcal{R}) \quad \text{bsup}_{\mathcal{P}(\mathcal{R})} \quad 1
\end{align*} \]

\[ \mathcal{P} \]

\[ \begin{cases} 
\text{QS (Quotient set)} \\
\text{AI (Abstract interpretation)} \\
+ \\
\end{cases} \rightarrow \text{formal approximation of quantitative properties} \]
Abstract Interpretation

Local Causality

\[
\text{local-cause}(d_1 \leadsto d_2) = \{ d_1 \xrightarrow{b_1} d_2, d_1 \xrightarrow{c_2} d_2 \}
\]

\[
\text{local-cause}^\#(d_1 \leadsto d_2) = \{ \{ b_1 \}, \{ c_2 \} \}
\]
**Initial context/state:** \( \langle a_0, b_0, c_1, d_1, e_1, f_0 \rangle \)

**Goal:** local state \( d_2 \)
Abstract Interpretation
Quantified Local Causality & Quotient Set

\[
\text{qlocal-cause}(d_1 \leadsto d_2) = \{ d_1 \xrightarrow{b_1,3} d_2, \\
d_1 \xrightarrow{c_2,2} d_2, \quad d_1 \xrightarrow{e_1,2} d_0 \} \\
\]

\[
\text{qlocal-cause}^\#(d_1 \leadsto d_2) = \{ \{ b_1 \}, \{ c_2 \}, \{ e_1 \} \} \\
\]

\[
\mathcal{P}_t(d_1 \leadsto d_2)(t) = \mathcal{P}(A_{d_1}, A_{d_2})(t) = \frac{R(A_{d_1}, A_{d_2})}{E(A_{d_1})} \cdot \mathcal{P}(A_{d_1}, *)(t) \\
\]

\[
\begin{array}{c}
\text{d}_2 \\
\{ \mathcal{P}(d_2), \mathcal{T}(d_2) \}
\end{array} \xrightarrow{ } \begin{array}{c}
\text{d}_1 \leadsto \text{d}_2 \\
\{ \mathcal{P}(d_1 \leadsto d_2), \mathcal{T}(d_1 \leadsto d_2) \}
\end{array} \xrightarrow{ } \begin{array}{c}
\text{e}_1 \\
\{ \mathcal{P}(e_1), \mathcal{T}(e_1) \}
\end{array} \\
\begin{array}{c}
\{ \mathcal{P}(b_1), \mathcal{T}(b_1) \}
\end{array} \xrightarrow{ } \begin{array}{c}
\text{b}_1 \\
\end{array} \xrightarrow{ } \begin{array}{c}
\{ \mathcal{P}(c_2), \mathcal{T}(c_2) \}
\end{array}
\]
Static analysis of quantitative properties in Stochastic Automata Networks

Quantitative Abstract Interpretation

Abstract Interpretation

Quantified Local Causality & Quotient Set

\[
\text{qlocal-cause}(d_1 \rightsquigarrow d_2) = \{ d_1 \xrightarrow{b_1,3} d_2, d_1 \xrightarrow{c_2,2} d_2, d_1 \xrightarrow{e_1,2} d_0 \}
\]

\[
\text{qlocal-cause}^\#(d_1 \rightsquigarrow d_2) = \{ \{b_1\}, \{c_2\}, \{e_1\} \}
\]

\[
\mathcal{P}_t(d_1 \rightsquigarrow d_2)(t) = \mathcal{P}_{(A_{d_1},A_{d_2})}(t) = \frac{R(A_{d_1},A_{d_2})}{E(A_{d_1})} \cdot \mathcal{P}_{(A_{d_1},*)}(t)
\]

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

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

Static analysis of quantitative properties in Stochastic Automata Networks

Probability/Delay

Louis FIPPO FITIME 35/48 LIPN seminar — 09/03/2017
Probability approximation

Static analysis of quantitative properties in Stochastic Automata Networks

Probability/Delay

Louis FIPPO FITIME 35/48 LIPN seminar — 09/03/2017
Probability approximation

Static analysis of quantitative properties in Stochastic Automata Networks

Probability/Delay
Probability approximation

Static analysis of quantitative properties in Stochastic Automata Networks — Probability/Delay

Probability approximation

- $d_1 \sim d_2$
- $b_0 \sim b_1$
- $c_1 \sim c_2$
- $a_0 \sim a_1$
- $f_0 \sim f_0$

Louis FIPPO FITIME
Probability approximation

Static analysis of quantitative properties in Stochastic Automata Networks

Proability/Delay

Louis FIPPO FITIME

35/48

LIPN seminar — 09/03/2017
Probability approximation

Static analysis of quantitative properties in Stochastic Automata Networks • Probability/Delay
Probability approximation

```
\[
\begin{align*}
&\text{Probability approximation} \\
&\quad \vdash \quad \text{Static analysis of quantitative properties in Stochastic Automata Networks} \\
&\quad \vdash \quad \text{Probability/Delay}
\end{align*}
\]

Louis FIPPO FITIME 35/48 LIPN seminar — 09/03/2017
Static analysis of bifurcations

Contributions

General properties: bounds
  nbr of attractors;
  functionality,...

Stable states

Bifurcations
  Reachability
    (CTL)

Quantitative reachability
  (CCSL)

Automata network

Chronology
Chronometry

1. A
Modeling

1. B
Simulation

Analysis

Simulation

Properties precision

Modelling

1
Static analysis

2
Static analysis

3
Static analysis

1
Simulation

2
Simulation

3
Simulation

Louis FIPPO FITIME 36/48 LIPN seminar — 09/03/2017
Static analysis of bifurcations

Contributions

General properties: bounds
number of attractors; functionality,...

Stable states

Bifurcations
Reachability (CTL)

Quantitative reachability (CCSL)

Chronometry

Automata network

Modeling

PINT

Simulation

Analysis

1.A

1.B

2

3

4
Static analysis of bifurcations

Motivation

stem cell differentiation

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

\[
\begin{align*}
E_F (g_1 \land \cdots) \\
\text{(CTL)}
\end{align*}
\]

\text{AI + LP} \quad \text{Approximation}

Logical Program (SAT/ASP)

\text{Approximation of bifurcations}

\text{Bifurcations PSPACE Complete}

\text{Contribution}: \quad \{ \begin{align*}
\text{LP (Logical Programming)} \\
+ \\
\text{AI (Abstract interpretation)}
\end{align*} \quad \Longrightarrow \text{formal approximation of bifurcations}
Static analysis of bifurcations

What is a bifurcation?

Illustration of bifurcations
Definition of bifurcation

$t_b$ 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)

**Interests:**
- Avoid state space explosion
- Decide in an efficient way reachability properties
Static analysis of bifurcations  

How to identify bifurcations?

Relaxation of bifurcation problem

# : sufficient conditions.
How to identify bifurcations?

Relaxation of bifurcation problem

\#: sufficient conditions.

(C1) \( s_b \cdot t_b \not\rightarrow^* g_1 \)  
(l1\#) \( \neg \text{OA}(s_u \rightarrow^* g_1) \)  
(l1\#) \( \Rightarrow \) (C1)
How to identify bifurcations?

Relaxation of bifurcation problem

#: sufficient conditions.

\[(C1) \ s_b \cdot t_b \not\rightarrow^* g_1 \quad (I1\#) \ \neg \ OA(s_u \rightarrow^* g_1) \quad (I1\#) \Rightarrow (C1)\]

\[(C2) \ s_b \rightarrow^* g_1 \quad (I2\#) \ \text{UA}(s_b \rightarrow^* g_1) \quad (I2\#) \Rightarrow (C2)\]
Relaxation of bifurcation problem

#: sufficient conditions.

(C1) \( s_b \cdot t_b \not\Rightarrow s_b \)  

(l1\#) \( \neg OA(s_u \Rightarrow s_b) \)  

(l1\#) \( \Rightarrow (C1) \)

(C2) \( s_b \Rightarrow s_b \)  

(l2\#) \( UA(s_b \Rightarrow s_b) \)  

(l2\#) \( \Rightarrow (C2) \)

(C3) \( s_0 \Rightarrow s_b \)  

(l3\#) \( UA(s_0 \Rightarrow s_b) \)  

(l3\#) \( \Rightarrow (C3) \)
Static analysis of bifurcations ○ How to identify bifurcations?

Relaxation of bifurcation problem

#: sufficient conditions.

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

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

(C3) $s_0 \rightarrow^* s_b$  
(I3) $s_b \in \text{reach}(s_0)$  
(I3#) $UA(s_0 \rightarrow^* s_b)$  
(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: $\text{poly}(\#\text{automata}), \text{exp}(|\text{single automaton}|)$

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
Static analysis of bifurcations ○ Biological applications

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

<table>
<thead>
<tr>
<th>Automata Network</th>
<th>Goal</th>
<th>M-C (NuSMV)</th>
<th>with (I3)</th>
<th>with (I3#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lambda phage</td>
<td>CI&lt;sub&gt;2&lt;/sub&gt;, Cro&lt;sub&gt;2&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>\Sigma</td>
<td>= 4$</td>
<td>$</td>
</tr>
<tr>
<td>EGF/TNF</td>
<td>NFkB&lt;sub&gt;0&lt;/sub&gt;, IKB&lt;sub&gt;1&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>\Sigma</td>
<td>= 28$</td>
<td>$</td>
</tr>
<tr>
<td>Th_th17</td>
<td>RORGT&lt;sub&gt;1&lt;/sub&gt;, BCL6&lt;sub&gt;1&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>\Sigma</td>
<td>= 101$</td>
<td>$</td>
</tr>
<tr>
<td>Th_HTG</td>
<td>BCL6&lt;sub&gt;1&lt;/sub&gt;, GATA3&lt;sub&gt;1&lt;/sub&gt;</td>
<td></td>
<td>out-of-time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>\Sigma</td>
<td>= 101$</td>
<td>$</td>
</tr>
</tbody>
</table>

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

Conclusions

Hybrid approach for modelling an RSTC network
- Formalizing biological knowledge: Translation of motifs in AN.
- Inference and integration of temporal parameters
- Qualitative and quantitative refinement of dynamics

Static analysis of quantitative property (Probability/delay)

\[
\begin{align*}
\text{QS (Quotient set)} & \quad + \\
\text{AI (Abstract interpretation)} & \quad \Rightarrow \quad \text{formal approximation of properties}
\end{align*}
\]

- Estimation of a lower bound of probability and delay of reachability property
- Good theoretical complexity

Identification of Bifurcations

\[
\begin{align*}
\text{AI (Artificial intelligence)} & \quad + \\
\text{AI (Abstract interpretation)} & \quad \Rightarrow \quad \text{formal approximation of bifurcations}
\end{align*}
\]

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

Stochastic simulation & statistic analysis
- Statistic analysis which take into account time.
- Use association rules to identify interesting relations between components.

Static analysis of probability
- Implementation (in progress...) of a static analyzer for properties in CCSL Logic
- Static analysis of long run behaviors (steady state)
- Static analysis of parametric models

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

Hybrid Modelling, Analysis and Verification of large-scale Biological Networks

Louis Fippo Fitime$^{1,3}$
 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)
$^3$ LCR / LIPN / CNRS (Paris, France)

olivier.roux@irccyn.ec-nantes.fr
 carito.guziolowski@irccyn.ec-nantes.fr
 loic.pauleve@lri.fr

Thank you for your attention!