

Learning Delays in Biological Regulatory Networks from Time Series Data



Emna Ben Abdallah¹

Tony Ribeiro¹, Morgan Magnin^{1,2}, Olivier Roux¹, Katsumi Inoue²

¹IRCCyN, École Centrale de Nantes, France, ²National Institute of Informatics, Tokyo, Japan

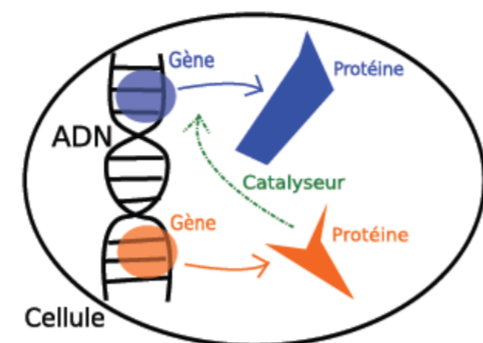


emna.ben-abdallah@irccyn.ec-nantes.fr, tony.ribeiro@irccyn.ec-nantes.fr, morgan.magnin@irccyn.ec-nantes.fr, olivier.roux@irccyn.ec-nantes.fr, inoue@nii.ac.jp

Introduction

- We want **all models**

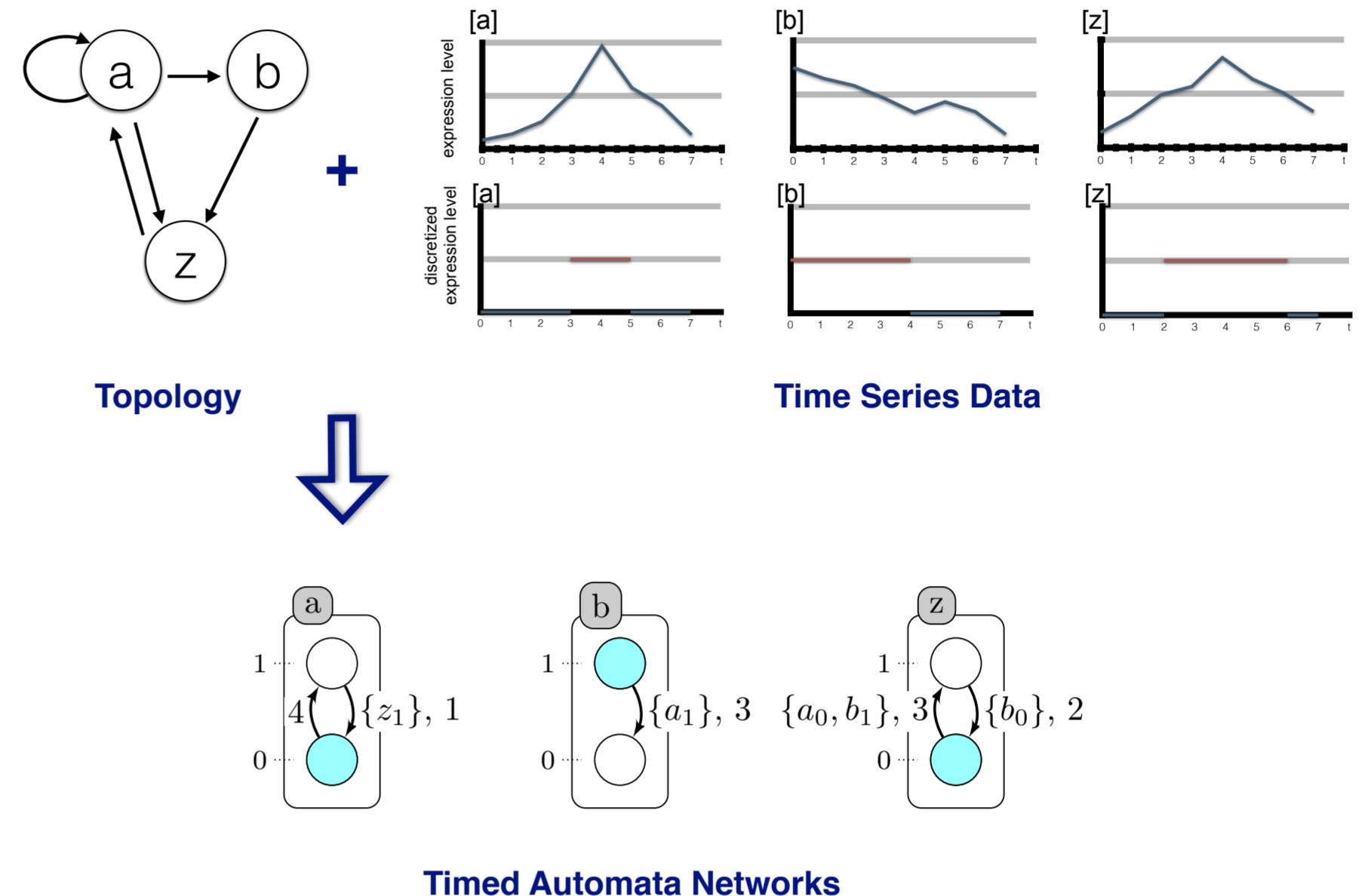
→ **consistent** with the topology;
→ can **reproduce** the time series data.



- Necessary conditions:

→ coherent thresholds and delays;
→ **realistic assumptions** for the model simulation;
→ non deterministic;
→ combined **asynchronous/ synchronous**.

Learning models

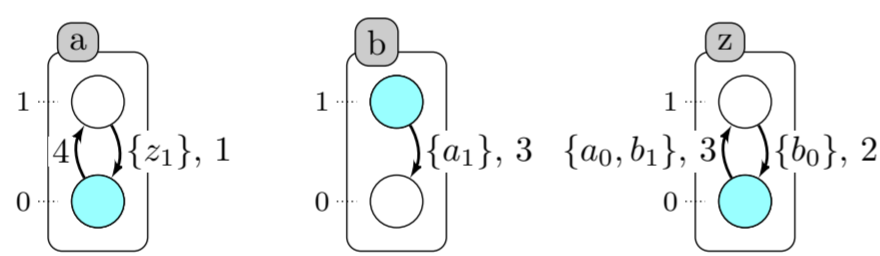


Modeling framework

Timed Automata Networks (T-AN):

- Well adapted to large-scale models [Paulevé in CMSB, 2016];
- qualitative levels for each component;
- atomistic representation of transitions;
- very efficient to analyze dynamical properties [Ben Abdallah et al. BIBM 2015].

Example:

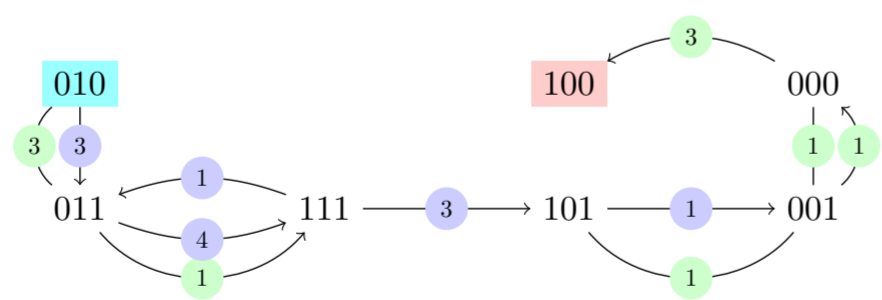


Dynamical semantics of T-AN:

Assumptions of our approach [Ben Abdallah et al. CMSB-2016]:

- allowing **simultaneity** between non conflicting transitions;
- blocking only** transitions in conflict (sharing same resources);
- memory** during the model simulation.

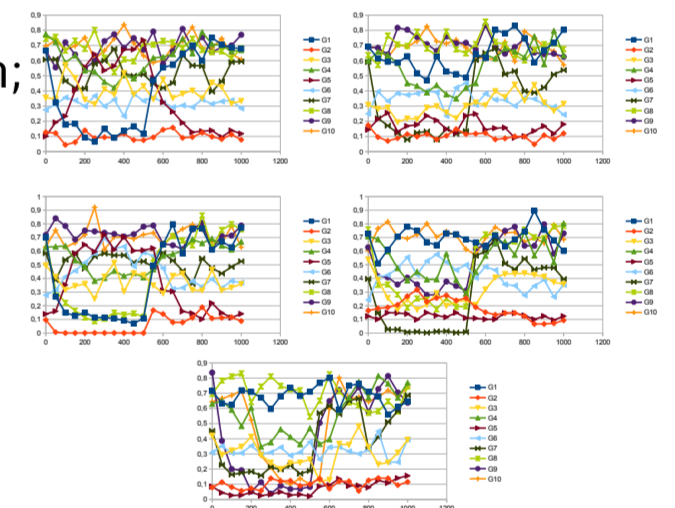
Example:



Application: DREAM4 Challenge

- DATA [Prill et al. Science signaling 2011]:

→ 5 different **time series** of genes expression;
→ each series has different perturbations;
→ topology (gold standard).



- GOAL:

→ model the network;
→ predict **thresholds** and **delays**.

- INPUT:

→ an initial state;
→ 5 different conditions of dual genes to be knockout simultaneously.

- OUTPUT:

→ predict the point attractor.

Evaluation

| Number of genes | Benchmarks | MSE |
|-----------------|------------|-------|
| 10 | 1 | 0.086 |
| | 2 | 0.080 |
| | 3 | 0.076 |
| | 4 | 0.039 |
| | 5 | 0.076 |
| 100 | 1 | 0.052 |
| | 2 | 0.042 |
| | 3 | 0.033 |
| | 4 | 0.033 |
| | 5 | 0.052 |

Mean Square Error (MSE):
difference between **predicted/expected** values.

Implemented using **Answer Set Programming**:
declarative approach & efficient solver.

Learning problematic

Main question:

what happened in **the past** that has led to the system state **dynamics**?

Main idea:

- detect biological components **changes**;
- compute the **candidate timed local transitions** responsible for the network changes (**threshold+delays**);
- generate **minimal subset** of candidate timed local transitions that can realize all changes;
- apply **filters** (contradiction, more frequent transitions, new knowledge...)

Conclusion & Perspectives

Work in progress:

- learning **circadian clock** model (perturbations: jet lag) and **DREAM11** challenge;
- identification of **basins of attraction** into large networks.

Perspectives:

- improving our implementation to make it **more robust** against noise;
- model checking** methods for Timed Automata Networks.