Modeling, Inference and Simulation of Biological Networks using Constraint Logic Programming

Fabien Corblin

IMAG, Université Joseph Fourier, France[†]
email: Fabien.Corblin@imag.fr

UNIVERSITE JOSEPH FOURIER
SCIENCES.TECHNOLOGIE.SANTÉ
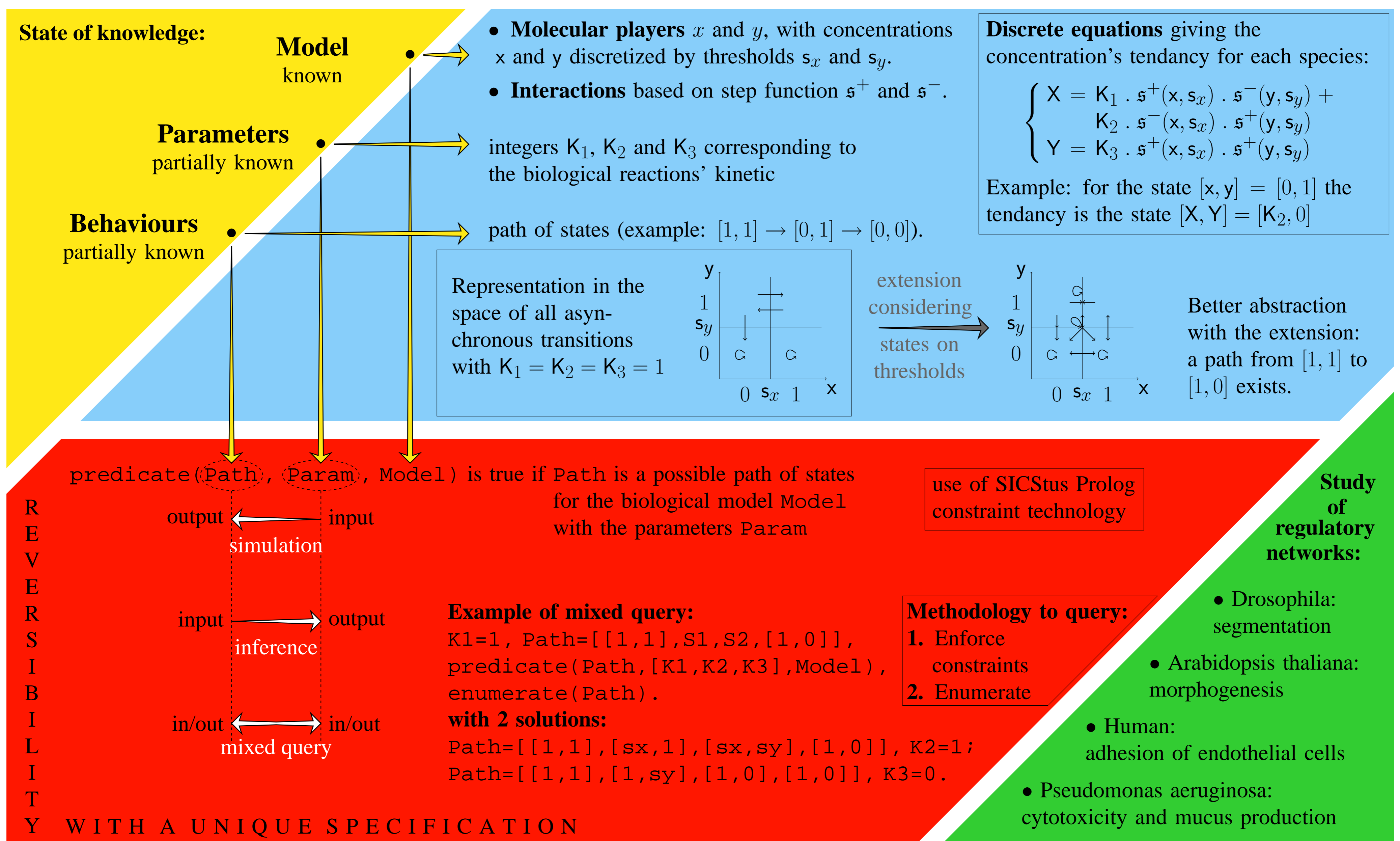GRENOBLE 1

Research summary

Computer tools are needed in systems biology to analyse qualitatively the dynamics of interaction networks, to discover the organisation of the cell's molecular components. In this context, the objective is to develop a general tool which, based on a unique specification, allows **to explore the properties of the model's parameters and behaviours**, and also to propose new experiments, by a mix of inference and simulation.

This work is based on the **multi-valued asynchronous networks** proposed by R. Thomas and E. Snoussi. This formalism, which can be seen as a discrete abstraction of a special class of piecewise affine differential equations, allows a qualitative analysis of the dynamic behaviour of such systems. This formalism has been recently extended by H. de Jong to take into account trajectories "sliding" along discretisation thresholds.

The goal of this research is to investigate how a formal description of such a biological switching network can be exploited through an implementation in **Constraint Logic Programming** (CLP) in order to obtain the variety of functionalities desired.

This tool will be applied to the **construction of several biological networks from observed properties**.

● Context ● , ● Formalism ● , ● Implementation ●   and   ● Applications ●

State of knowledge:

Model known

Parameters partially known

Behaviours partially known

- **Molecular players** x and y, with concentrations x and y discretized by thresholds s_x and s_y.
- **Interactions** based on step function $\mathfrak{s}^+$ and $\mathfrak{s}^-$.

integers K_1, K_2 and K_3 corresponding to the biological reactions' kinetic

path of states (example: $[1,1] \to [0,1] \to [0,0]$).

Discrete equations giving the concentration's tendency for each species:

$$\begin{cases} X = K_1 . \mathfrak{s}^+(x,s_x) . \mathfrak{s}^-(y,s_y) + \\ \qquad K_2 . \mathfrak{s}^-(x,s_x) . \mathfrak{s}^+(y,s_y) \\ Y = K_3 . \mathfrak{s}^+(x,s_x) . \mathfrak{s}^+(y,s_y) \end{cases}$$

Example: for the state $[x,y] = [0,1]$ the tendancy is the state $[X,Y] = [K_2,0]$

Representation in the space of all asynchronous transitions with $K_1 = K_2 = K_3 = 1$

extension considering states on thresholds

Better abstraction with the extension: a path from $[1,1]$ to $[1,0]$ exists.

R E V E R S I B I L I T Y

`predicate(Path, Param, Model)` is true if `Path` is a possible path of states for the biological model `Model` with the parameters `Param`

output ← input
simulation

input → output
inference

in/out ← → in/out
mixed query

use of SICStus Prolog constraint technology

Example of mixed query:
```
K1=1, Path=[[1,1],S1,S2,[1,0]],
predicate(Path,[K1,K2,K3],Model),
enumerate(Path).
```
with 2 solutions:
```
Path=[[1,1],[sx,1],[sx,sy],[1,0]], K2=1;
Path=[[1,1],[1,sy],[1,0],[1,0]], K3=0.
```

Methodology to query:
1. Enforce constraints
2. Enumerate

Study of regulatory networks:

- Drosophila: segmentation
- Arabidopsis thaliana: morphogenesis
- Human: adhesion of endothelial cells
- Pseudomonas aeruginosa: cytotoxicity and mucus production

W I T H   A   U N I Q U E   S P E C I F I C A T I O N

Perspectives

Efficiency: by using the structure of the discrete equations to enhance the tool (current work).

Extension of models: in order to be more accurate regarding biology and so to obtain a correct abstraction.

Interface: in order to provide a usable tool for a biologist (not trivial in the context).