A formal model for gene regulatory networks with time delays

Jean-Paul Comet[1], Jonathan Fromentin[2], Gilles Bernot[1], and Olivier Roux[3]

[1] Laboratoire I3S, UMR 6070 UNS-CNRS, Université de Nice
2000, route des Lucioles, B.P. 121, 06903 Sophia Antipolis CEDEX, France
{bernot,comet}@unice.fr
[2] Labri, Université de Bordeaux, 33 Talence, France
jonathan.fromentin@labri.fr
[3] IRCCyN UMR 6597, CNRS & École Centrale de Nantes
1, rue de la Noë - BP 92 101 - 44321 Nantes CEDEX 03 - France
olivier.roux@irccyn.ec-nantes.fr

Abstract. We introduce a hybrid modelling framework for gene regulatory networks as an extension of the René Thomas' discrete modelling framework. We handle temporal aspects through *delays* expressing the time mandatory to pass from a qualitative state to another one. It permits one to build, from a specification expressed in terms of paths, the constraints on the temporal parameters in order to assure the consistency between the hybrid model and the specification.
We illustrate this modelling framework on the simple system of *mucus* production in the bacterium *Pseudomonas aeruginosa*. We show through this example how to build the constraints on the delays parameters for the specification of a cycle in the dynamics.

1 Introduction

Modelling gene regulatory networks aims at deep understanding of their behaviours and thus at some non-obvious predictions [1–4]. Unfortunately, while available data on the interaction graph between genes are more and more numerous, the kinetic data allowing us to identify the sensible parameters are difficult to obtain experimentally. This parameter identification problem constitutes the cornerstone of the modelling activities. Whereas the quantitive models (differential equations, stochastic models) need a good precision on the available information about the dynamics of the system, qualitative models which focus only on the qualitative features of the dynamics, make easier the parameter identification problem. This comment motivates the development of different methods for which this identification problem is tractable [5–7]. For example René Thomas' discrete modelling [8] of gene regulatory networks (GRN) is a well-known approach to study the dynamics resulting from a set of interacting genes. It deals with some *discrete* parameters that reflect the possible targets of trajectories. Those parameters are *a priori* unknown, but they can generally be deduced from a well-chosen set of biologically observed trajectories. Moreover

there exists a strong correspondence between modelling by piecewise linear differential equations and such boolean or discrete modellings [9].
Unfortunately this framework neglects the time delay necessary for a gene to pass from one level of expression to another one, whereas information on the time necessary for the system to go from one state to another one is often experimentally available. For example, time spent by the system to cover a whole turn of a periodic trajectory (*e.g.* circadian cycle) is often well known. Such kind of information is not used to face up the parameter identification problem in the "standard" Thomas' framework without delays. This remark motivated several researchers to develop mathematical frameworks [10] or formal frameworks [11–15] where time is explicit. The effect of time delays on the robustness of differential systems becomes also an interesting research perspective [16].

In this article, we propose a new modelling framework which extends the discrete modelling framework of René Thomas by introducing temporal aspects. This modelling framework inherits from the pure qualitative modelling framework the computer aided methods for determination of suitable parameter values, but it introduces a continuous notion of time through the handling of delays. Thus, we propose a hybrid modelling framework where discrete and (temporal) continuous dynamics are mixed. Naturally, these delays are coded by new parameters, *i.e.* delays mandatory for a gene to go from a discrete abstract level to another one, which are not deductible from previous qualitative models. When this framework is viewed as an abstraction of piecewise linear differential equations, as discrete modelling is, some constraints on the delays of the hybrid model can be built to ensure the consistency between the hybrid model and the underlying system of piecewise linear differential equations. In particular, delays of the hybrid models have to satisfy some constraints which can be deduced from the piecewise linear differential equation systems. Nevertheless, kinetic parameters of the PLDE system are generally unknown and the key point of the modelling process lies in identification of these kinetic parameters. Adding delays, the identification problem is more difficult because of the increased number of parameters. Nonetheless much temporal data is available from experiments and because hybrid modelling frameworks preserve powerful computer-aided reasoning capabilities, computer is able to reject a large class of parameter values.
To illustrate our hybrid modelling framework, we use as *running example*, an extremely simplified model, representing the production of *mucus* of the bacterium *Pseudomonas aeruginosa* [17, 18]. *P. aeruginosa* is an opportunistic pathogen, often encountered in chronic lung diseases such as cystic fibrosis. The main regulator for the mucus production, AlgU, supervises an operon which is made of 4 genes among which one codes for a protein that is an inhibitor of AlgU. Moreover AlgU favours its own synthesis. The mucus production regulatory network can then be simplified into a regulatory graph with two nodes: x represents AlgU, and y its inhibitor [17]. x regulates positively y and also regulates itself, whereas y regulates negatively x. From a biological point of view, it is crucial to determine if the change of behaviours (passing from a state where mucus is not produced to another one where it is) is mostly due to change of the regulations

(mutation) or mostly due to a change of state. We show in this article that it is possible to construct a hybrid model in which the behaviour which does not produce mucus is represented by a limit cycle.

The paper is organized as follows. We first recall in section 2 the principles of the modelling by a system of piecewise linear differential equations and of the discrete modelling of René Thomas. Section 3 is devoted to the definition of the considered hybrid models. In section 4, we sketch how to build a set of constraints on the delays parameters in order to get a hybrid model whose dynamics present a particular path. Finally section 5 is devoted to concluding remarks.

2 Continuous and discrete models

PLDE modelling. Modelling a gene regulatory network with a system of piecewise linear differential equations [19], PLDE for short, makes mandatory the knowledge of regulations. In particular, for each regulation, which can involve several regulators, one has to define under which real concentration conditions this regulation is effective. As usual, because regulations are often considered as sigmoidal, we consider only the piecewise differential system, which is built as an approximation of the differential system by replacing sigmoids by steps functions: $s^+_\theta(x) = \begin{cases} 1, & x > \theta \\ 0, & x < \theta \end{cases}$ and $s^-_\theta(x) = 1 - s^+_\theta(x)$ where $\theta \in \mathbb{R}^+$ is the threshold of the sigmoid.

Definition 1 (PLDE). *Let us consider a finite set of positive real variables $X = \{x_1, x_2, \ldots, x_n\}$ and let us denote x the vector $(x_1, x_2, \ldots, x_n)$. A system of piecewise linear differential equations (PLDE) on X is defined by:*

$$\dot{x}_i = g_i(x) - \gamma_i x_i \quad with \quad 0 \leq x_i \quad and \quad 1 \leq i \leq n$$

where γ_i is the degradation rate of variable x_i and each g_i is a function representing the synthesis rate of variable x_i which is supposed to be additive (the synthesis rate is the sum of all effective regulations):

$$g_i(x) = k_i + \sum_{j \in \mathcal{R}(i)} k_{ij} r_{ij}(x) \tag{1}$$

where

- *$k_i \in \mathbb{R}^+$ and $k_{ij} \in \mathbb{R}^{+*}$ are kinetic parameters,*
- *The regulation functions r_{ij} are some combinations of step functions:*

$$< r > ::= s^+_\theta \,|\, s^-_\theta \,|\, 1 - < r > \,|\, < r > \times < r >$$

- *$\mathcal{R}(i)$ is the set of possible indices such that r_{ij} is a regulation function on i.*

The dynamics of a PLDE system is intrinsically related to kinetic parameters. In the rest of the paper, kinetic parameters are indexed by a *set of resources*. Intuitively, the set of resources at a given continuous state is the set of the regulations which are effective at this continuous state.

Definition 2 (Resources). *The set of resources of variable x_i at continuous state x, denoted $\Omega_i(x)$, is the finite set $\Omega_i(x) = \{j \mid r_{ij}(x) = 1\}$.*

Because of the finite number of possible sets of resources, the concentration space of each variable x_i can be partitioned in equivalence classes defined by the same set of resources of variable x_i: x_i^1 and x_i^2 are in the same equivalence class iff $\Omega_i(x_i^1) = \Omega_i(x_i^2)$. These equivalence classes split the concentration space of x_i into open intervals which can be classicaly numbered by 0, 1... : 0 is the *name* of the first interval, 1 denotes the second interval and so on.

We extend this equivalence relation to the concentration space of n dimensions. The principle of the partition is simple: we gather in the same *domain* all the continuous states for which each concentration coordinate is in the same interval. Because of the form of the regulation functions, all domains (*i.e.* equivalence classes) are hyper-rectangular zones. Moreover, since all the continuous states of the same domain are identically situated with regard to the thresholds, they all have the same set of resources: $\forall x \in d$, $\Omega_i(x)$ =constant. So we can define the set of resources of a domain:

Definition 3 (Resources of a domain). *The set of resources of variable x_i in domain d, denoted $\omega_i(d)$ is the set of resources (see Def. 2) of variable x_i at any point x of d: $\omega_i(d) = \{j \mid \forall x \in d, r_{ij}(x) = 1\}$.*

Finally, to simulate a PLDE system, values of kinetic parameters (k_i and k_{ij} in eq. (1)) have to be given. Unfortunately, these parameters are not easy to evaluate *in vivo*, and values obtained *in vitro* are not necessarily transposable for the system *in vivo*. Valuating parameters thus becomes the cornerstone of the modelling process.

Discrete modelling. To overcome these difficulties of parameters valuation, René Thomas first introduced a boolean framework [7] then a discrete formalism [8] which have been proven to be consistent with the PLDE modelling framework [9]. In this section, we sketch this qualitative framework which mimics qualitativelly the continuous framework.

From a qualitative point of view, at a particular point of the concentration space, the dynamics is controlled only by the set of the regulations which are resources. Actually René Thomas did not propose such a rich way to describe the regulations but this discrete modelling framework can be easely extended. Let us first notice, that the regulations do not change inside a same domain class, that is, the differential equation system is linear in each hyper-rectangular zone which define the domains. Then the solutions in each zone are analytically deducible and converge towards a unique *focal point*. Then

- with each domain is associated a qualitative state,
- the coordinate i of the focal point associated to the domain d is given by $((k_i + \sum_{j \in \mathcal{R}(i)} k_{ij} r_{ij}(x))/\lambda_i)_{i \in V}$ for any $x \in d$,
- because of the monotonicity of the solutions of the differential equations, trajectories starting in the domain d go towards the associated focal point until they reach the boundary of d.

– From a qualitative point of view, only the position of the focal point is important. Then, for the domain d we call $K_{i,\omega_i(d)}$ the number of the interval in which stays the coordinate i of the focal point which depends only on the set of resources $\omega_i(d)$ of variable x_i in domain d.

This idea leads to the definition of the discrete transition system.

Definition 4 (Transition system). *The discrete dynamics of a gene regulatory network with n variables is given by the* transition system *defined by:*

– *the set of vertices is the set of equivalence classes of the concentration space, called a discrete states; each equivalence class is represented by a vector of integer $d = (d_i)_{i\in[1,n]}$ where d_i is the number of the interval in which stays the coordinate i of a particular point of the equivalence class,*
– *There exists a transition from the discrete state d to the discrete state d' if*
  - $\exists i \in [1,n]$ *such that* $\begin{cases} d'_i = d_i + 1 & \text{and} \quad K_{i,\omega_i(d)} > d_i \\ d'_i = d_i - 1 & \text{and} \quad K_{i,\omega_i(d)} < d_i \end{cases}$
  - $\forall j \neq i, d'_j = d_j$.

A strategy for determining discrete parameters. Let us observe that the number of different parameters in the discrete modelling framework is finite and that each parameter can take a finite number of values. Thus, by enumaration, all the possible models can be simulated in order to keep only the valuations of parameters leading to a transition system which is consistent with all the available specifications on the behaviour of the biological system. Generally, known behavioural properties can be expressed by a particular qualitative observation of the following class: the saturation of the cell in a particular gene product (resp. the knock-out of a gene) leads to a state where an other specific gene product is present or absent.

This computer aided modelling approach has already been implemented using classical model-checking techniques [18] or symbolic model-checking techniques [20], and then using constraint programming techniques [21]. The observation data are transcripted into temporal logic formulas, a formal representation of a knowledge about the traces of a system which can be handled by computers. In [18], for each possible valuation, the transition system is computed and a procedure of model-checking is performed. This allows one to retain only the valuations that lead to a transition system satisfying the formula. This approach, requiring enumeration of all parameter valuations, has been rephrased for a temporal logic so that a single pass of model-checking gives a symbolic representation of all the models validating the temporal property [20]. The approaches adopted in [21–23] use constraints programming. The temporal logic formula is translated into constraints on the discrete parameters of the model. These constraints also symbolically represent all the parameter valuations that lead to transition systems satisfying the formula.

3 Hybrid Modelling

Because real time is ignored in the complete discrete modelling framework, some qualitative behaviours are not distinguishable. For example, an inward spiral is abstracted by the same discrete model than a outward spiral. This remark motivated us to introduce a modelling framework that combines the discrete modelling framework with temporal delays while preserving consistency with PLDE systems.

Syntactical features of hybrid models. We associate with each domain a *temporal zone* which measures the time elapsed in the domain. This zone is represented as a n-dimensional hypercube (where n is the number of variables in the system) whose edges have various lengths. Intuitively, the length of the hypercube in the i-axis represents the mandatory delay for the system to entirely cross the associated domain (in concentration) along the i-axis.

Definition 5 (State graph with delays (SGD)). *Let $G = (g_i(x_i))_{x_i \in X}$ be the regulation schema, which defines the synthesis rate of each variable according to the effectiveness of each regulation (see equation 1). A* State Graph with Delays *(SGD for short) associated with the regulation schema G is a 4-tuple $\mathcal{N} = (X, L, K, D)$ where:*

- *$X = \{x_1, \ldots, x_n\}$ is the set of variables,*
- *$L = \{(l_i(x_i))_{x_i \in X}\}$ is the finite set of domains deduced from G by the equivalence relation on the concentration space; for each $x \in X$, we define the integer b_x as the number of different thresholds describing the different actions of x on its targets,*
- *$K = \{K_{x,\omega}\}_{x \in X, \omega \subset \mathcal{R}(x)}$ is a family of integers such that $K_{x,\omega} \in [0, b_x]$ for any variable x and for any set ω of regulations on x.*
- *$D = D^+ \cup D^-$ is a family of positive real numbers such that:*
  - *$D^+ = \{\delta^+_{x,i,\omega}\}_{x \in X, i \in [0,b_x], \omega \subset \mathcal{R}(x), i \leq K_{x,\omega}}$ with $\delta^+_{x,i,\omega} \in \mathbb{R}^+$ ($[0, b_x]$ being an interval of integers).*
  - *$D^- = \{\delta^-_{x,i,\omega}\}_{x \in X, i \in [0,b_x], \omega \subset \mathcal{R}(x), i \geq K_{x,\omega}}$ with $\delta^-_{x,i,\omega} \in \mathbb{R}^+$ ($[0, b_x]$ being an interval of integers).*

  The subfamily D^+ is called the set of production delays of $\mathcal{N}$ and the subfamily D^- is called the set of degradation delays of $\mathcal{N}$.

Intuitively, for a given domain d, the temporal zone is defined by the product of intervals: $\prod_{x \in X} [0, \delta^+_{x,l(x),\omega_x(d)} + \delta^-_{x,l(x),\omega_x(d)}]$

Running example. Let us now consider the SGD $\mathcal{P} = (X, L, K, D)$ modelling the system of mucus production of *Pseudomonas Aeruginosa*, defined by:

- $X = \{x, y\}$,
- $L = \{(0,0), (1,0), (0,1), (1,1), (2,0), (2,1)\}$,
- $K = \{K_{x,\emptyset} = 0, K_{x,\{x\}} = 2, K_{x,\{y\}} = 2, K_{x,\{x,y\}} = 2, K_{y,\emptyset} = 0, K_{y,\{x\}} = 1\}$

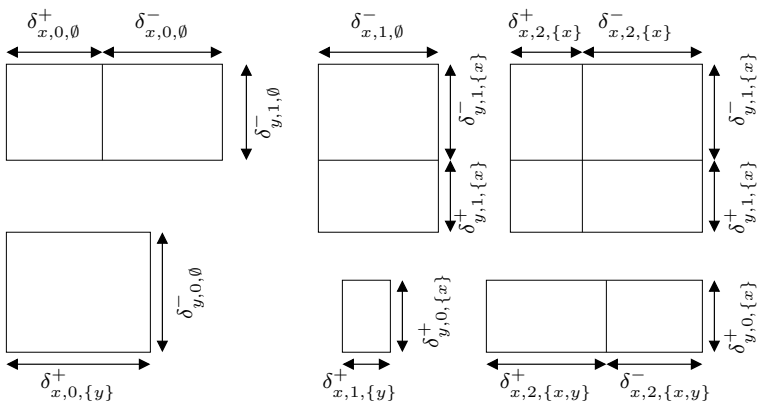

Fig. 1. State graph with delays modelling the system of the mucus production by *Pseudomonas aeruginosa*. Only non-zero delays are drawn.

- $D = D^+ \cup D^-$ where
  - $D^+ = \{\delta^+_{x,2,\{x\}}, \delta^+_{x,2,\{x,y\}}, \delta^+_{x,1,\{y\}}, \delta^+_{x,0,\{y\}}, \delta^+_{y,1,\{x\}}, \delta^+_{y,0,\{x\}}\}$ and
  - $D^- = \{\delta^-_{x,2,\{x\}}, \delta^-_{x,2,\{x,y\}}, \delta^-_{x,1,\emptyset}, \delta^-_{x,0,\emptyset}, \delta^-_{y,1,\emptyset}, \delta^-_{y,0,\emptyset}, \delta^-_{y,1,\{x\}}\}$.

This *state graph with delays* is represented in Figure 1.

Semantics of hybrid models: the dynamics. Let us observe that to specify a particular state of a SGD, one needs a couple of values: the first value is a domain, and the second is a point in the associated temporal zone. More formally, for a given SGD $\mathcal{N} = (X, L, K, D)$, a *state* of $\mathcal{N}$ is a couple $\eta = (l, \tau)$ where:

- $l : X \to \mathbb{N}$ is a domain of $\mathcal{N}$ (i.e. $l \in L$).
- $\tau : X \to \mathbb{R}^+$ is a total function s.t. $\forall v \in X, \tau(v) \leq \delta^+_{v,l(v),\omega_v(l)} + \delta^-_{v,l(v),\omega_v(l)}$
  The real number $\tau(v)$ is called the *delay residue* of v at the level $l(v)$.

As we already mentioned, temporal zones allow one to measure the time elapsed in a domain. Intuitivelly, the evolution in the model is twofold :

- inside a domain, the point in the temporal zone evolves in a *linear* way, it measures the time spent in a domain along a given evolution direction.
- to pass from a domain l to another one, it is mandatory that the point in the temporal zone reaches a border. If the point reaches the face for which the delay residue $\tau(v)$ is null (resp. equal to $\delta^+_{v,l(v),\omega_v(l)} + \delta^-_{v,l(v),\omega_v(l)}$), the system leaves the previous domain and enters into the new domain where the concentration level $l(v)$ is decremented (resp. incremented). The face of the temporal zone that is reached defines the new (accessible) domain.

To go further in the formalization of these ideas, we introduce two kinds of delays. The first one is the mandatory time for a variable to allow the system to move from a domain to another one: it is the *moving delay*, see figure 2. Unfortunatelly, this definition is not sufficient to determine if the reached face allows the exit from the domain. Thus, the *cross delay* is introduced.

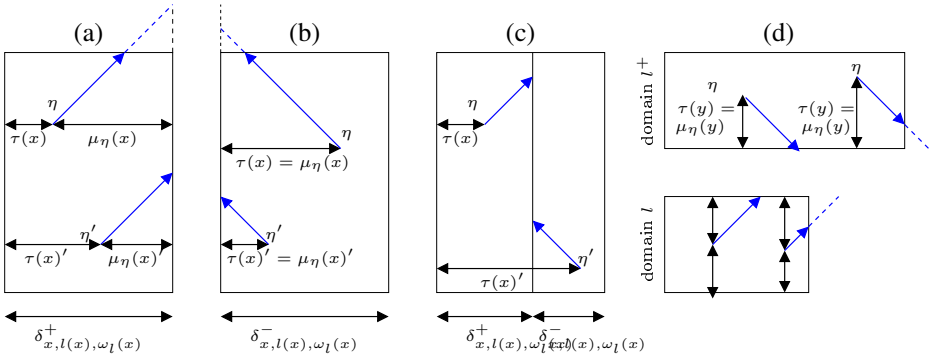

Fig. 2. Moving and cross delays. **(a)** When $l(x) < K_{x,\omega_l(x)}$ the moving delay is $\delta^+_{x,l(x),\omega_l(x)} - \tau(x)$ in the x-direction (horizontal). **(b)** When $l(x) > K_{x,\omega_l(x)}$ the moving delay is $\tau(x)$ in the x-direction (horizontal). **(c)** When $l(x) = K_{x,\omega_l(x)}$, x cannot be responsible for the exit from the domain. Thus the moving delay is $\mu_\eta(x) = \infty$. **(d)** Illustration of cross delays when $\mu_\eta(y) \neq \infty$ and $\bar{\mu}_\eta(y) = \infty$ (y-axis is the vertical one).

Definition 6 (moving & cross delays). *Let $\eta = (l, \tau)$ be a state of a* SGD $\mathcal{N}$*,*

- *the* moving delay *of a variable v is given by the function $\mu_\eta : X \to \mathbb{R}^+ \cup \{\infty\}$ defined by* $\mu_\eta(v) = \begin{cases} \infty & \text{if } K_{v,\omega_v(l)} = l(v) \\ |\delta^+_{v,l(v),\omega_v(l)} - \tau(v)| & \text{if } K_{v,\omega_v(l)} \neq l(v) \end{cases}$.
- *the* cross delay *of a variable v is given by the function $\bar{\mu}_\eta : X \to \mathbb{R}^+ \cup \{\infty\}$ defined by:*
  - *If ($K_{v,\omega_v(l)} < l(v)$ and $K_{v,\omega_v(l^-)} > l(v) - 1$) or ($K_{v,\omega_v(l)} > l(v)$ and $K_{v,\omega_v(l^+)} < l(v) + 1$) then $\bar{\mu}_\eta(v) = \infty$,*
  - *else $\bar{\mu}_\eta(v) = \mu_\eta(v)$.*

  where domains l^+ and l^- are such that $\forall u \neq v$, $l^+(u) = l^-(u) = l(u)$ and $l^+(v) - 1 = l^-(v) + 1 = l(v)$.

The moving delay of variable v is simply the time necessary for this variable to allow the system to exit from the current domain. If variable v is not able to reach the boundary of the temporal zone, the moving delay of variable v is ∞, see Fig. 2-(c). When $\mu(v) \neq \infty$, the cross delay of variable v can nevertheless be equal to ∞ when v is attracted outside the current domain, but cannot exit in that direction since, beyond the limit of the domain, this variable is immediately attracted again inside the domain. This stands for the notion of sliding modes [19]. Illustration of such a situation is given in Fig. 2-(d).

The temporal evolutions from a state within the temporal zone are linear: directions of these evolutions are given by the following definition.

Definition 7 (Discrete partial derivative). *Given a domain l of a* SGD $\mathcal{N}$*, for any state $\eta = (l, \tau)$ and for any variable v, the* discrete partial derivative *of $\mathcal{N}$ at l with respect to v, $\kappa_l(v)$, is defined by:*

– *if* $l(v) < K_{v,\omega_l(v)}$ *and* $\bar{\mu}_\eta(v) \neq \infty$ *then* $\kappa_l(v) = 1$
– *if* $\bar{\mu}_\eta(v) = \infty$ *then* $\kappa_l(v) = 0$
– *if* $l(v) > K_{v,\omega_l(v)}$ *and* $\bar{\mu}_\eta(v) \neq \infty$ *then* $\kappa_l(v) = -1$

We can now define the successor states of a state (see Fig 3) using the function *sign*: $sign(x) = 1$ if $x > 0$, $sign(x) = -1$ if $x < 0$ and $sign(x) = 0$ if $x = 0$.

Definition 8 (Successor). *A state* $\eta' = (l', \tau')$ *of a* SGD $\mathcal{N}$ *is a* successor *state of the state* $\eta = (l, \tau)$ *if there exists a variable* $x \in X$ *such that:*

1. $\forall y \in X,\ \bar{\mu}_\eta(x) \leq \bar{\mu}_\eta(y)$,
2. $l'(x) = l(x) + \kappa_l(x)$,
3. $\forall y \in X,\ y \neq x \Rightarrow l'(y) = l(y)$,
4. $\kappa_l(x) = 1 \Rightarrow \tau'(x) = 0$,
5. $\kappa_l(x) = -1 \Rightarrow \tau'(x) = \delta^+_{x,l'(x),\omega_x(l')} + \delta^-_{x,l'(x),\omega_x(l')}$,
6. $\forall y \in X$ *such that* $y \neq x$ *and* $\kappa_l(y) \neq 0$,
   $\kappa_l(x) \neq 0 \Rightarrow \tau'(y) = \frac{\left(\tau(y) + \text{sign}(\delta^+_{y,l(y),\omega_y(l)} - \tau(y)) \times \mu_\eta(x)\right) \times \left(\delta^+_{y,l'(y),\omega_y(l')} + \delta^-_{y,l'(y),\omega_y(l')}\right)}{\delta^+_{y,l(y),\omega_y(l)} + \delta^-_{y,l(y),\omega_y(l)}}$,
7. $\forall y \in X$ *such that* $y \neq x$ *and* $\kappa_l(y) = 0$,
   $\kappa_l(x) \neq 0 \Rightarrow \tau'(y) = \frac{\left(\tau(y) + \text{sign}(\delta^+_{y,l(y),\omega_y(l)} - \tau(y)) \times \min(\mu_\eta(x), \mu_\eta(y))\right) \times \left(\delta^+_{y,l'(y),\omega_y(l')} + \delta^-_{y,l'(y),\omega_y(l')}\right)}{\delta^+_{y,l(y),\omega_y(l)} + \delta^-_{y,l(y),\omega_y(l)}}$,
8. $\kappa_l(x) = 0 \Rightarrow \left(\forall y \in X,\ \tau'(y) = \delta^+_{y,l(y),\omega_y(l)}\right)$.

If $\kappa_l(x) \neq 0$, *then the* transition time *from* η *to* η' *is* $\zeta(\eta, \eta') = \mu_\eta(x)$. *If* $\kappa_l(x) = 0$, *then* $\zeta(\eta, \eta')$ *is equal to* $\min_{v \in X}(\mu_\eta(v))$.

Note that in item 6 of the previous definition, the computation of new delay residue for variable y depends on the sign of $(\delta^+_{y,l(y),\omega_l(y)} - \tau(y))$. Indeed, if $\delta^+_{y,l(y),\omega_l(y)} < \tau(y)$ (resp. $\delta^+_{y,l(y),\omega_l(y)} > \tau(y)$), the coordinate $\tau(y)$ decreases (resp. increases) towards $\delta^+_{y,l(y),\omega_l(y)}$.
The previous definition covers both of the following cases.

1. Let us first focus on the case where a domain contains its focal point (see Fig. 3-b). Temporal trajectories do not go out of this domain: all the cross delays are equal to ∞, and each discrete partial derivative is null. Thus, we can take for x any element of X (see item 1). Items 2 and 3 imply that $l' = l$. Finally item 8 gives the temporal coordinates of the focal point. Transition time is then the time necessary for each variable y to reach the coordinate f_y of the focal point.
2. We now focus on a domain which does not contain its focal point (see Fig. 3-a). Each temporal trajectory goes out of this domain passing a threshold on one v-axis. This variable v is the one which has the smallest *not-infinite* cross delay (see item 1). Items 2 and 3 imply that l' differs from l on only one coordinate. Items 4 and 5 reset *residue delay* associated with v whereas items 6 and 7 compute the new *residue delays* associated with the other variables (these expressions come from the homothetic transformation). The transition time is then the time to reach the face of the temporal zone, that is the moving delay.

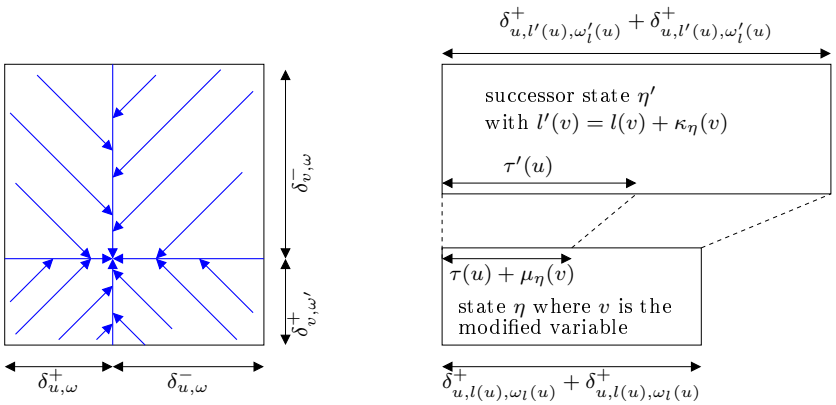

Fig. 3. Illustration of Definition 8. **(a)** the domain contains its focal point and all the cross delays are infinite. **(b)** the domain does not contain its focal point.

Definition 9 (State space). *The* state space *of a* RND $\mathcal{N}$ *is the (infinite) directed "graph" $S_\mathcal{N}$ the vertices of which are the states of $\mathcal{N}$ and the edges of which are the couples (η, η') such that η' is a successor of η. Given a path $p = \eta_0\eta_2\cdots\eta_n$, the* crossing time *of p is defined as $\tau(p) = \sum_{i=1}^{n} \tau(\eta_{i-1}, \eta_i)$.*

Transitions between domains are the transitions of the discrete dynamics in the formalism of René Thomas.

4 Construction of the Delays Constraints

The parameter values of the hybrid model can be straightforwardly deduced from a PLDE system with known parameters. For the discrete part, the parameters correspond to the position of the steady states of the linear differential system of the considered domain, whereas the parameters of family D correspond to the time mandatory for the differential system to cross the domain. Let us just mention that when the v-coordinate of the focal point f of the domain l is inside $l(v)$ then, neither the production delay nor the degradation delay is null: $\delta^+_{v,k,\omega(l)}$ (resp. $\delta^+_{v,k,\omega(l)}$) measures the duration between the time when a trajectory gets into the domain by the face having the smaller (resp. the bigger) v-concentration value and the time when the coordinate v of the focal point f is reached. Whereas from the point of view of PLDE, this time is infinite, for the hybrid model this time is not infinite.

But in general, when modelling a biological regulatory network, we have only a partial knowledge about the form of the regulatory functions (r_{ij}). Specifically, kinetic parameters of the PLDE system are unknown and the key point of the modelling process thus lies in identification of these kinetic parameters. Paragraph about *strategies for determining discrete parameters* of section 2 sketches, in the context of purely discrete modelling, a computer aided method for helping

in this task. In our context of hybrid modelling, even if the qualitative parameters ($K_{x,\omega}$) are assumed to be known (or deduced from a computer aided approach), it remains to determine which values of the time delays are actually consistent with known properties of the studied system.

Once again, we start from some knowledge about the dynamics of the studied biological system. This knowledge often comes from experimental observations which are expressed as paths in the discrete transition system. These paths constitute the *specifications* since it determines the set of models which have to be considered. This section sketches how these specifications build up the models, and more accurately a system of parameter constraints.

The principle of the construction of these constraints relies on the enumeration of constraints due to paths of length 2: $\mu_0 \rightarrow \mu_1 \rightarrow \mu_2$. For a longer path, the constraint is the conjunction of constraints due to each sub-path of length 2.

For sake of readability, we describe here only one situation among twelve[1]. Let us consider the path $\mu_0 \rightarrow \mu_1 \rightarrow \mu_2$ where the first (resp. second) transition is due to a qualitative increasing of variable i_0 (resp. i_1). Let us suppose moreover that the vector $(c_i)_{i \in V}$ represents the delays residue when entering into μ_1 and that there exists in μ_1 a variable i'_1 which can also increase. In order to allow the global path $\mu_0 \rightarrow \mu_1 \rightarrow \mu_2$, the following relation has to be satisfied:

$$(d^+_{i_1}(\mu_1) - c_{i_1}) < (d^+_{i'_1}(\mu_1) - c_{i'_1})$$

Processing discrete cycle. The discrete cycles can abstract several different behaviours: fully cyclic temporal trajectories, convergent spirals, divergent spirals, limit cycle, etc. Thus, it is interesting to know more precisely their behaviours in the hybrid modelling. For example, it can be proved that the discrete cycle of *Pseudomonas aeruginosa* – $(0,0) \rightarrow (1,0) \rightarrow (1,1) \rightarrow (0,1) \rightarrow (0,0)$ – can abstract different kinds of qualitative behaviours of hybrid models. In other words, from the same purely discrete model with a discrete cycle, it is possible to construct a hybrid model which presents either: (1) a set of convergent spirals or (2) a set of cyclic temporal trajectories which constitute a torus and that we call fully cyclic temporal trajectories or (3) a set of divergent spirals or (4) a limit cycle, that is, a torus of volume null (see Fig. 4).

5 Conclusion

We developed a new hybrid modelling framework for gene regulatory networks which extends the discrete modelling framework of René Thomas by introducing temporal features through delays handling. These delays express the time mandatory to pass from a qualitative state to another.

On the one hand, this modelling framework inherits from the differential modelling framework, since it is possible to build an hybrid model consistent with the underlying system of piecewise linear differential equations (PLDE).

[1] The other cases are addressed in a similar enough way and the proof can be sent upon request.

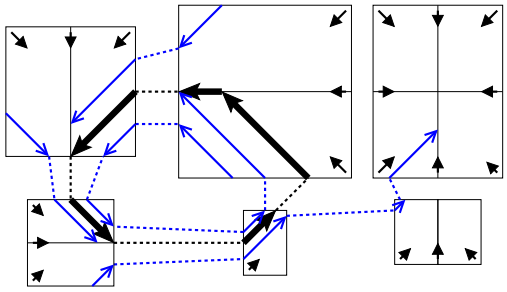

Fig. 4. A particular hybrid dynamics with a limit cycle (in thick black line) modelling the system of the mucus production by Pseudomonas aeruginosa.

On the other hand, this modelling framework inherits also from the pure qualitative modelling framework, the computer aided methods for determination of suitable parameter values, but it introduces a continuous notion of time through delays handling. When kinetic parameters are not available, it is possible to build some constraints on the new delays parameters in order to get a model satisfying a specification expressed in terms of paths. Finally, adding information about delays in the qualitative framework allows one to distinguish qualitatively different behaviours which are abstracted into a common purely discrete model.

With such hybrid frameworks, systems biology should take advantage of the whole corpus of formal methods from computer science which opens a large horizon of research perspectives. It will be necessary to develop for example algorithms that compute the set of parameter valuations that are compatible with reachability properties. Indeed, hybrid modellings are not the ultimate aim, they are only a guideline for predictions that suggest biological experiments, whose success will be *in fine* the discriminent criterion. In such a perspective, hybrid approaches could constitute a trade-off between expressiveness and computational tractability.

References

1. Ideker, T., Galitski, T., Hood, L.: A new approach to decoding life: systems biology. Annual Rev. Genomics Hum. Genet. **2** (2001) 343–372
2. Oltvai, Z., Barabási, A.: Systems biology. Life's complexity pyramid. Science **298**(5594) (2002) 763–764
3. Kitano, H.: Computational systems biology. Nature **420**(6912) (2002) 206–210
4. Conti, F., Valerio, M., Zbilut, J., Giuliani, A.: Will systems biology offer new holistic paradigms to life sciences? Syst Synth. Biol. **1**(4) (2007) 161–165
5. Rashevsky, N.: Mathematical Biophysics: Physico-Mathematical Foundations of Biology. University of Chicago Press (1948)
6. Sugita, M.: Functional analysis of chemical systems in vivo using a logical circuit equivalent. Journal of Theoretical Biology **1** (1961) 415–430
7. Thomas, R.: Boolean formalization of genetic control circuits. Journal of Theoretical Biology **42** (1973) 563–585

8. Thomas, R.: Regulatory networks seen as asynchronous automata : A logical description. Journal of Theoretical Biology **153** (1991) 1–23
9. Snoussi, E.: Qualitative dynamics of a piecewise-linear differential equations : a discrete mapping approach. Dynamics and stability of Systems **4** (1989) 189–207
10. Farcot, E., Gouzé, J.L.: Limit cycles in piecewise-affine gene network models with multiple interaction loops. International Journal of Systems Science **41**(1) (2010) 119–130
11. Siebert, H., Bockmayr, A.: Incorporating time delays into the logical analysis of gene regulatory networks. In: CMSB. Volume 4210 of LNCS. (2006) 169–183
12. Ahmad, J., Bernot, G., Comet, J.P., Lime, D., Roux, O.: Hybrid modelling and dynamical analysis of gene regulatory networks with delays. ComPlexUs **3**(4) (2007) 231–251
13. Batt, G., Ben Salah, R., Maler, O.: On timed models of gene networks. In: FORMATS. Volume 4763 of LNCS., Springer (2007) 38–52
14. Maler, O., Pnueli, A.: Timing analysis of asynchronous circuits using timed automata. In: in Proc. CHARME'95, LNCS 987, Springer (1995) 189–205
15. Comet, J.P., Bernot, G.: Introducing continuous time in discrete models of gene regulatory networks. In: Proc. of the Nice Spring school on Modelling and simulation of biological processes in the context of genomics. EDP Sciences, ISBN : 978-2-7598-0545-7 (2010) 61–94
16. Radde, N.: The impact of time-delays on the robustness of biological oscillators and the effect of bifurcations on the inverse problem. Eurasip J. Bioinf. Syst. Biol. (2009)
17. Guespin-Michel, J., Kaufman, M.: Positive feedback circuits and adaptive regulations in bacteria. Acta. Biotheor. **49** (2001) 207–218
18. Bernot, G., Comet, J.P., Richard, A., Guespin, J.: Application of formal methods to biological regulatory networks: Extending Thomas' asynchronous logical approach with temporal logic. Journal of Theoretical Biology **229**(3) (2004) 339–347
19. de Jong, H., Gouzé, J.L., Hernandez, C., Page, M., Sari, T., Geiselmann, J.: Qualitative simulation of genetic regulatory networks using piecewise-linear models. Bull. Math. Biol. **66**(2) (2004) 301–40
20. Mateus, D., Gallois, J.P., Comet, J.P., Le Gall, P.: Symbolic modeling of genetic regulatory networks. J. of Bioinformatics and Comput. Biol. **5**(2B) (2007) 627–640
21. Fromentin, J., Comet, J.P., Le Gall, P., Roux, O.: Analysing gene regulatory networks by both constraint programming and model-checking. In: EMBC07, 29th IEEE EMBS Annual Intern. Conf., IEEE Press (2007) 4595–4598
22. Fanchon, E., Corblin, F., Trilling, L., Hermant, B., Gulino, D.: Modeling the molecular network controlling adhesion between human endothelial cells: Inference and simulation using constraint logic programming. In: CMSB, Springer (2004) 104–118
23. Corblin, F., Fanchon, E., Trilling, L.: Modélisation de réseaux biologiques discrets en programmation logique par contraintes. Technique et Science Informatiques **26**(1-2) (2007) 73–98