Comparing Discrete and Piecewise Affine Differential Equation Models of Gene Regulatory Networks

Shahrad Jamshidi, Heike Siebert, and Alexander Bockmayr

Freie Universität Berlin, DFG Research Center Matheon
Arnimallee 6, D-14195 Berlin, Germany
{jamshidi,hsiebert,bockmayr}@zedat.fu-berlin.de

Abstract. We compare the discrete asynchronous logical modeling formalism for gene regulatory networks due to R. Thomas with piecewise-affine differential equation models. We show that although the two approaches are based on equivalent information, the resulting qualitative dynamics are different.

Key words: Gene regulatory networks, mathematical modeling, discrete models, piecewise-affine models

1 Introduction

Gene regulation is the result of the complex interplay of molecular components forming large interaction networks. Mathematical modeling of gene regulatory networks gives insights into the underlying structure and dynamics of various biological systems. If information on kinetic parameters is lacking, *qualitative formalisms* offer a well-established alternative to the more traditionally used differential equation models. Using only qualitative information on the network structure and the interactions between the components, these approaches allow obtaining an abstract description of the system's dynamics.

The discrete formalism of R. Thomas [1] is a qualitative method describing a gene regulatory network by a discrete function. Each network component is represented by a variable that takes integer values representing the different levels of gene activity. The information on how the behavior of one component is governed by the values of the other components is captured in a discrete function. The component functions then constitute the coordinate functions of the update function of the network. To derive the dynamics of the system, Thomas introduced the asynchronous update method where only one variable changes per discrete time step, and only by a unit value. Since the state space is finite, the dynamics can be represented by a directed graph, the so-called asynchronous *state transition graph* (STG).

The particularities of the asynchronous update method result in a close correspondence of the discrete model to certain differential equation systems [2].

Differential equation models using step functions have a continuous time evolution, yet can be seen as qualitative due to the close relation of step and discrete functions. Such piecewise affine differential equation (PADE) models approximate certain ordinary differential equation models [3, 4]. De Jong et al. [5] have shown that they can essentially be captured by a discrete representation which abstracts the continuous solution trajectories of the differential equations into transitions between different regions of the phase space. Again, the resulting dynamics can be represented by a digraph, the *qualitative transition graph* (QTG).

In this paper, we aim at clarifying the relation between Thomas and PADE models by comparing the respective graphs capturing the dynamical behavior. Several results in this direction already exist. For example, attractors, including steady states and certain limit cycles, are related [2, 3, 6–8]. Our goal here is to present a comprehensive comparison between the STG and the QTG.

The paper is organized as follows. Sect. 2 presents a discrete modeling approach based on the Thomas formalism. PADE systems and the qualitative analysis developed by de Jong et al. [5] are introduced in Sect. 3. In Sect. 4, we show that the two approaches use equivalent information. Sect. 5 contains our main result characterizing transitions in the QTG using edges originating in corresponding vertices in the STG. We illustrate the application of this result with examples of relations between paths and attractors in the two graphs. The conclusion and perspectives for future work are given in Sect. 6.

This paper is an extended abstract intended to illustrate more formal results.

2 Discrete formalism

Consider a gene network with n regulatory components. In the discrete modeling approach, the activity level of a component i is modeled by a discrete variable q_i, which takes its values in a finite set of natural numbers $Q_i = \{0, \ldots, p_i\}$. The *state space* of the discrete model is $Q = Q_1 \times \cdots \times Q_n$, and the regulatory interactions are captured by a discrete *update function* $f = (f_1, \ldots, f_n) : Q \to Q$. The function f uniquely determines the *state transition graph* $STG(f) = (Q, E)$, a directed graph with node set Q and edge set $E \subset Q \times Q$. For any $j \in \{1, \ldots, n\}$, $q \in Q$ with $f_j(q) \neq q_j$, there is an edge $(q, q') \in E$, where $q'_j = q_j + sgn(f_j(q) - q_j)$ and $q'_i = q_i$, for all $i \in \{1, \ldots, n\} \setminus \{j\}$. Here, $sgn : \mathbb{R} \to \{-1, 0, 1\}$ denotes the sign function. If $f(q) = q$, then $(q, q) \in E$ and q is called a *fixpoint*.

Example 1. For $Q = \{0, 1\} \times \{0, 1, 2\}$ and the update function $f : Q \to Q$

q	00	01	02	10	11	12
$f(q)$	12	12	11	00	10	11

the state transition graph $STG(f)$ is displayed on the left of Fig. 1(a).

3 Piecewise affine differential equations

Next we discuss piecewise affine differential equations (PADE) and the qualitative modeling approach introduced by de Jong et al. [5].

Consider an n-dimensional phase space $\Omega = \Omega_1 \times \cdots \times \Omega_n \subset \mathbb{R}^n_{\geq 0}$, where $\Omega_i = \{x_i \in \mathbb{R} \mid 0 \leq x_i \leq \max_i\}$, and $\max_i \in \mathbb{R}_{>0}$. For every continuous variable $x_i \in \Omega_i$ we assume $p_i \in \mathbb{N}$ thresholds $\theta_i^1, \cdots, \theta_i^{p_i}$ satisfying the ordering

$$0 < \theta_i^1 < \cdots < \theta_i^{p_i} < \max_i, \quad \text{for all } i \in \{1, \ldots, n\}. \tag{1}$$

In the comparison with the discrete formalism in Sect. 2, the value p_i chosen here corresponds to the maximal value p_i of the component range Q_i of a discrete model. We consider a set of PADEs in Ω of the form

$$\dot{x}_i = F_i(x) - G_i(x)x_i, \quad i \in \{1, \ldots, n\}, \tag{2}$$

where the functions $G_i : \Omega \to \mathbb{R}_{>0}$ and $F_i : \Omega \to \mathbb{R}_{\geq 0}$ are linear combinations of products of step functions $S^+(x_l, \theta_l^k) = \begin{cases} 0 & \text{if } x_l < \theta_l^k, \\ 1 & \text{if } x_l > \theta_l^k, \end{cases}$ and $S^-(x_l, \theta_l^k) = 1 - S^+(x_l, \theta_l^k)$ for $l \in \{1, \ldots, n\}$.

To obtain a discrete representation of the PADE system, the state space is partitioned into a set of domains.

Definition 1. *Consider a set of PADEs of the form (2) with phase space Ω and thresholds θ_i^j. The $(n-1)$-dimensional hyperplanes corresponding to the equations $x_i = \theta_i^j$, $j \in \{1, \ldots, p_i\}$, divide Ω into hyper-rectangular regions called* domains*. A domain $D \subset \Omega$ is defined by $D = D_1 \times \cdots \times D_n$ where every D_i is given by one of the following equations*

$$\begin{aligned}
&D_i = \{x_i \,|\, 0 \leq x_i < \theta_i^1\}, \\
&D_i = \{x_i \,|\, \theta_i^k < x_i < \theta_i^{k+1}\} \text{ for } k \in \{1, \ldots, p_i - 1\}, \\
&D_i = \{x_i \,|\, \theta_i^{p_i} < x_i \leq \max_i\}, \\
&D_i = \{x_i \,|\, x_i = \theta_i^k\} \text{ for } k \in \{1, \ldots, p_i\}.
\end{aligned}$$

By $\mathcal{D}$ we denote the set of all domains in Ω. A domain $D \in \mathcal{D}$ is called a singular domain*, if there exists $i \in \{1, \ldots, n\}$ such that $D_i = \{x_i \,|\, x_i = \theta_i^k\}$ for some $k \in \{1, \ldots, p_i\}$. The variable x_i is then called* singular variable*. The* order *of a singular domain is the number of its singular variables. A domain $D \in \mathcal{D}$ is called a* regular domain*, if it is not a singular domain. The set of regular and singular domains are denoted by $\mathcal{D}_r$ and $\mathcal{D}_s$ respectively.*

It follows immediately that for any regular domain $D \in \mathcal{D}_r$, the functions $F_i(x)$ and $G_i(x)$ are constant on D. Thus (2) can be written as a linear system $\dot{x} = F^D - G^D x$, for all $x \in D$, where $G^D = diag(G_1^D, \ldots, G_n^D)$ is a diagonal matrix with strictly positive entries and $F^D = (F_1^D, \ldots, F_n^D)$ a positive vector. It is easy to see that solutions of (2) starting in a regular domain D converge monotonically towards the so-called *focal point* $\phi(D) := (G^D)^{-1} F^D$.

In agreement with [5], we assume that all focal points lie in a regular domain. By definition of the regular domains, we can then encode the position of each focal point by strict inequalities using the threshold values, and thus obtain a parameter constraint of the form (1) consisting of threshold values and components of the focal point. We call these constraints *ordering constraints*.

To define a suitable dynamics of (2) on singular domains, the differential equations are extended to differential inclusions, and methods presented in [5, 9, 10], give us so-called Fillipov solutions of the differential inclusions. However, our focus here is on the qualitative dynamics, which does not depend on the particularities of the Fillipov extension.

It is shown in [5] that we can calculate the qualitative dynamics of a PADE system (2) using only the respective ordering constraints. This dynamics is represented by a directed graph, the *qualitative transition graph* $QTG(\mathcal{A}) = (\mathcal{D}, \mathcal{T})$. Here, the node set $\mathcal{D}$ consists of all regular and singular domains, and the arcs indicate the existence of suitable solution trajectories of (2) between adjacent domains. As shown in [5], all systems in the class of PADE systems satisfying the same ordering constraints have the *same* qualitative dynamics, i.e., isomorphic QTGs.

Example 2. Consider the system of PADEs

$$\begin{aligned}\dot{x}_1 &= \alpha_1 + \beta_1 S^+(x_1, \theta_1^1) S^-(x_2, \theta_2^1) - \lambda_1 x_1,\\ \dot{x}_2 &= \alpha_2 + \beta_2 S^+(x_1, \theta_1^1) S^-(x_2, \theta_2^2) + \gamma_2 S^-(x_1, \theta_1^1) S^-(x_2, \theta_2^2) - \lambda_2 x_2.\end{aligned}$$

The system has six regular domains with corresponding focal points, e.g., the focal point of $D = [0, \theta_1^1) \times [0, \theta_2^1)$ being $(\frac{\alpha_1}{\lambda_1}, \frac{\alpha_2+\gamma_2}{\lambda_2})$. We impose the ordering constraints $0 < \frac{\alpha_1+\beta_1}{\lambda_1} < \theta_1^1 < \frac{\alpha_1}{\lambda_1} < \max_1$ and $0 < \frac{\alpha_2+\beta_2}{\lambda_2} < \theta_2^1 < \frac{\alpha_2}{\lambda_2} < \theta_2^2 < \frac{\alpha_2+\gamma_2}{\lambda_2} < \max_2$. The resulting QTG is given on the right of Fig. 1(a).

4 Relating the discrete and the PADE formalism

Now we show that the PADE and the discrete formalism contain the same information in the sense that we can transform a PADE system with given ordering constraints into a discrete update function and vice versa.

To obtain a discrete update function from a PADE system we can use a straightforward method originally proposed by Snoussi [2]. First, we discretize the continuous phase space of the PADE system according to its threshold values.

Definition 2. *Let $\mathcal{A}$ be a set of PADEs as in (2), where each variable x_i has p_i ordered threshold values. Let $Q := Q_1 \times \cdots \times Q_n$, where $Q_i := \{0, 1, \ldots, p_i\}$, $i \in \{1, \ldots, n\}$. Define the bijective mapping $d^{\mathcal{A}} : \mathcal{D}_r \to Q$, where*

$$d_i^{\mathcal{A}}(D) := \begin{cases} 0 & \text{if } D_i = \{x \in \mathbb{R} \mid 0 \le x < \theta_i^1\},\\ q & \text{if } D_i = \{x \in \mathbb{R} \mid \theta_i^q < x < \theta_i^{q+1}\},\\ p_i & \text{if } D_i = \{x \in \mathbb{R} \mid \theta_i^{p_i} < x \le \max_i\}.\end{cases}$$

Second, we exploit the localization of the focal points in the regular domains in order to construct an update function $f^{\mathcal{A}} : Q \to Q$ on the discretized state space Q that shares the dynamical properties of the PADE system $\mathcal{A}$ (see [2] for details). The function $f^{\mathcal{A}}$ is uniquely determined by the ordering constraints for $\mathcal{A}$. Consequently, the set of PADE systems $\mathcal{A}$ satisfying given ordering constraints can be associated with a single discrete update function $f^{\mathcal{A}}$.

Conversely, a discrete update function can easily be transformed into a PADE system that shares the qualitative dynamical properties.

Definition 3. *Let $f : Q \to Q$ be an update function. Define the discretization function d corresponding to the thresholds $\theta_j^k = k - \frac{1}{2}$ for $j \in \{1, \ldots, n\}$, $k \in \{1, \ldots, p_j\}$, according to Def. 2. We denote by $PADE(f)$ the system of PADEs on $\Omega := \prod_{i=1}^{n} [0, \max_i]$, $\max_i \in \mathbb{R}_{>p_i}$ for all $i \in \{1, \ldots, n\}$, of the form*

$$\dot{x} = F(x) - x_i, \quad \textit{where} \quad F(x) = \sum_{q \in Q} f(q) \prod_{j=1}^{n} S(x, q)$$

and S is composed componentwise of products of step functions yielding $S(x, q) = 1$ if $x \in d^{-1}(q)$, and $S(x, q) = 0$ otherwise.

The choice of threshold values is generic, ensuring an obvious correspondence between the values $0, 1, \ldots, p_i$ in Q_i, $i \in \{1, \ldots, n\}$, and the intervals $[0, \theta_i^1)$, $(\theta_i^k, \theta_i^{k+1})$ for $k \in \{1, \ldots, p_i - 1\}$, and $(\theta_i^{p_i}, \max_i]$. If we calculate the regular domains according to the threshold values and their focal points, we have $\phi(D) = F(x) = f(d(D))$ for all $x \in D$, where $d := d^{PADE(f)}$ and $D \in \mathcal{D}_r$. Equivalently, it holds that $\phi(d^{-1}(q)) \in d^{-1}(f(q))$ for $q \in Q$, which immediately implies that the focal points of $PADE(f)$ satisfy the set of corresponding ordering constraints.

Using these two transformations, we can associate a class of PADE systems characterized by their ordering constraints with a unique discrete update function, and vice versa. The information necessary for constructing the STG resp. QTG is inherent in both representations. In that sense, we can identify every STG with a QTG and vice versa. In fact, it can easily be shown that this identification can be achieved solely on the level of the graphs representing the respective dynamics, although the graphs may hold less specific information than the corresponding discrete functions resp. PADE systems. That is, already the information inherent in the STG is enough to construct the corresponding QTG and vice versa.

Example 3. The function f from Ex. 1 generates $PADE(f)$ whose parameter values satisfy the ordering constraints of the PADE system $\mathcal{A}$ from Ex. 2. Thus $PADE(f)$ belongs to the PADE class represented by $\mathcal{A}$. Similarly, if we discretize $\mathcal{A}$ using Snoussi's method, we obtain the update function f from Ex. 1.

5 Comparing the dynamics

Although the STG and QTG can be obtained from each other, the qualitative dynamics represented by the two graphs is not the same. Next we analyze differences and similarities of $STG(f)$ and $QTG(\mathcal{A})$ for a discrete update function $f : Q \to Q$ and the corresponding PADE system $\mathcal{A} := PADE(f)$.

Initially, we compare the node sets. The discretization in Sect. 4 implies that the vertices of $STG(f)$ correspond to the regular domain vertices of $QTG(\mathcal{A})$. However, there is no representation of the singular domains in the purely discrete

setting. To overcome this problem we associate with every singular domain D the set $H(D) \subset Q$ corresponding to the discretization of those regular domains D' that have D in their *boundary* $\partial D'$ (cf. [5]). We thus introduce the mapping

$$H(D) := \begin{cases} \{d(D)\}, & \text{if } D \in \mathcal{D}_r, \\ \{d(D') \in Q \mid D \subset \partial D', D' \in \mathcal{D}_r\} & \text{if } D \in \mathcal{D}_s. \end{cases}$$

Now we are able to state our main result on the correspondences between edges in $QTG(\mathcal{A}) = (\mathcal{D}, \mathcal{T})$ and $STG(f) = (Q, E)$.

Theorem 1. *Let $D \in \mathcal{D}$ and $D' \subset \partial D$. Denote by I the index set of singular variables in D and I' the index set of singular variables in D'. Then*

1. *$(D, D') \in \mathcal{T}$ if and only if*
   (a) *for all $i \in I$ there exist $q^1, q^2 \in H(D)$, $q^1 \neq q^2$, such that $p_i^1 \leq q_i^1$, $q_i^2 \leq p_i^2$ for all $p^1, p^2 \in Q$ with $(q^1, p^1), (q^2, p^2) \in E$, and*

$$\exists\, l \in \{1,2\},\, p \in H(D):\ (q^l, p) \in E \ \wedge\ p_i \neq q_i^l, \ \text{ if } \ q_i^1 = q_i^2,$$
$$\exists\, p^1, p^2 \in H(D):\ (q^1, p^1), (q^2, p^2) \in E \ \wedge\ p_i^1 < p_i^2, \ \text{ if } \ q_i^1 > q_i^2,$$

   (b) *for all $i \in I' \setminus I$ there exists $q \in H(D)$ and $q' \in H(D') \setminus H(D)$ with $q_i \neq q_i'$ and $(q, q') \in E$,*
2. *$(D', D) \in \mathcal{T}$ if and only if*
   (a) *condition 1.(a) holds, and*
   (b) *for all $i \in I' \setminus I$ there exists $q \in H(D)$ and $q' \in H(D') \setminus H(D)$ such that $q_i \neq q_i'$, $q_j' = q_j$ for all $j \neq i$ and $(q, q') \notin E$.*

This result provides the basis for elucidating the correspondences between more complex structures, such as paths or attractors. On the one hand, it can be used for proofs building on local considerations concerning the edges involved and thus confirming some previous findings [7,10]. On the other hand, it provides ideas for the construction of counterexamples, some of which we present here to illustrate that the relation between the two dynamics is not clear-cut.

We start by considering reachability properties. In simple cases, we can find conditions ensuring the existence of corresponding paths. However, reachability properties are not conserved in general, as can be seen from Fig. 1 (a). There, state (0,2) is reachable from (1,0) in the STG via the path indicated in gray. In the QTG, all paths starting in the regular domain corresponding to (1,0) and all adjacent singular domains do not cross the first threshold plane of the second component. In Fig. 1 (b), we see by considering paths from (0,0) to (1,1) that reachability properties of the QTG are also not conserved in the STG.

Similarly, we are able to find correspondences of attractors (terminal strongly connected components of the directed graphs) including certain steady states, i.e., singleton attractors, and limit cycles, in accordance with [2,3,6–8]. However, the situation becomes less clear if we consider more general attractors.

First, we consider the number of attractors. In Fig. 1 (a) both systems have one attractor, but the STG in Fig. 1 (b) has two (a fixed point and a cyclic

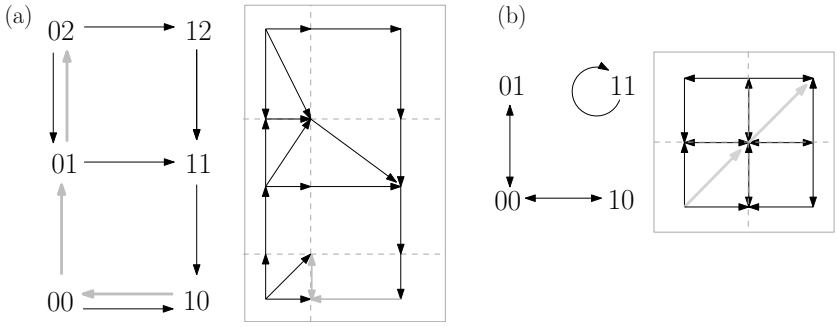

Fig. 1. Corresponding STGs and QTGs. The partitioned phase space of a corresponding PADE underlying a QTG is shown in fine gray lines (dashed for threshold planes) underneath the QTG and allows identification of nodes representing singular resp. regular domains. In (a), STG of Ex. 1 and QTG of Ex. 2. In (b), the STG of a two component Boolean network with the corresponding QTG below. Heavier gray edges illustrate reachability properties discussed in the text.

attractor) while the QTG has only one (a steady state in the upper right node). In Fig. 2 (a), the STG has fewer attractors than the corresponding QTG.

In addition, the relation between the attractor structure is not clear-cut. While the cyclic attractor of the STG in Fig. 1 (a) comprises all nodes and contains nodes with multiple outgoing edges, the cyclic attractor in the QTG is a simple cycle consisting only of two nodes joined by the heavier gray double edge in the lower part of the graph. In Fig. 1 (b) the cyclic attractor in the STG vanishes in the corresponding QTG. The same happens in Fig. 2 (b), but here an additional steady state can be observed in a singular node.

These examples illustrate that, in general, neither the number nor the character of the attractors is preserved. It can be shown that hyper-rectangular trap sets, i.e., node sets that no path can leave, correspond in the two graphs. This may be helpful in further elucidating the correspondences of attractors.

6 Discussion and perspectives

In summary, the information inherent in the STG of the discrete update function is sufficient to derive the QTG of the corresponding PADE system and vice versa. Despite this fact, many characteristics of the two graphs are not preserved. This implies that, contrary to what might be expected, the QTG of the PADE system is not a straightforward refinement of the STG of the Thomas model.

Motivated by these findings, there are several directions for future work. First, we would like to better understand and characterize the network properties that lead to substantial differences, e.g., in the number of attractors in the dynamics of the Thomas and the PADE model. Second, we plan to extend the analysis to closely related formalisms like the refined qualitative representation of PADE

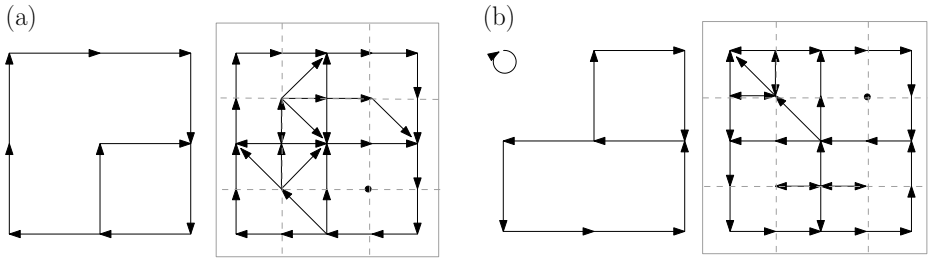

Fig. 2. Two examples for networks with two components and three activity levels for each component. In each case, the STG is depicted on the left, the corresponding QTG on the right. Depiction of the graphs corresponds to that in Fig. 1, only the explicit labeling of the STG nodes is omitted. In (a), the STG consists of a single cyclic attractor, while the QTG has an additional steady state at the lower right singular node depicted by a fat dot. In (b), both STG and QTG have a steady state in the upper left node. The STG has an additional cyclic attractor. The QTG has no cyclic attractor, but a singular steady state at the upper right singular node depicted by a fat dot.

systems [11] as well as piecewise multi-affine models [12]. Also, there exist approaches that allow the integration of threshold values directly into the Thomas formalism [13, 14]. Clarifying the relation between the different approaches may allow one to transfer available results and analysis methods from one formalism to the other. Also, progress in this direction may be helpful when deciding on the most suitable and efficient modeling framework in a concrete application.

References

1. Thomas, R. and D'Ari, R. *Biological Feedback*. CRC Press, (1990).
2. Snoussi, E. H. *Dyn. Syst.* **4**(3), 565–583 (1989).
3. Glass, L. and Kauffman, S. *J. Theoret. Biol.* **34**(2), 219–237 (1972).
4. Glass, L. and Kauffman, S. A. *J. Theoret. Biol.* **39**(1), 103–129 (1973).
5. de Jong, H., Gouzé, J.-L., Hernandez, C., Page, M., Sari, T., and Geiselmann, J. *Bull. Math. Biol.* **66**(2), 301–40 (2004).
6. Snoussi, E. H. and Thomas, R. *Bull. Math. Biol.* **55**(5), 973–991 (1993).
7. Chaves, M., Tournier, L., and Gouzé, J.-L. *Acta Biotheoretica* **58**(2), 217–232 (2010).
8. Wittmann, D. M., Krumsiek, J., Saez-Rodriguez, J., Lauffenburger, D. A., Klamt, S., and Theis, F. J. *BMC Systems Biology* **3**, 98 (2009).
9. Fillipov, A. *Differential equations with discontinuous righthand sides*. Springer, (1988).
10. Gouzé, J.-L. and Sari, T. *Dyn. Syst.* **17**(4), 299–316 (2002).
11. Batt, G., de Jong, H., Page, M., and Geiselmann, J. *Automatica* **44**(4), 982–989 (2008).
12. Kloetzer, M. and Belta, C. *Transactions of the Institute of Measurement and Control* **32**(5), 445–467 (2009).
13. Thomas, R., Thieffry, D., and Kaufman, M. *Bull. Math. Biol.* **2**(57), 24–276 (1995).
14. Richard, A., Bernot, G., and Comet, J.-P. *Fund. Inform.* **65**, 373–392 (2005).