Graphic requirements for multistability and attractive cycles in a Boolean dynamical framework

Élisabeth Remy*, Paul Ruet*, Denis Thieffry† [1]

*CNRS - Institut de Mathématiques de Luminy, UMR 6206
Campus de Luminy, Case 907, 13288 Marseille Cedex 9 (France).
Email: {remy,ruet}@iml.univ-mrs.fr

†Technologies Avancées pour le Génome et la Clinique, UMR 628, INSERM, Université de la Méditerranée, Campus de Luminy, Case 928, 13288 Marseille Cedex 9 (France). Email: thieffry@tagc.univ-mrs.fr

Abstract

To each Boolean function $f : \{0,1\}^n \to \{0,1\}^n$ and each $x \in \{0,1\}^n$, we associate a signed directed graph $G(x)$, and we show that the existence of a positive circuit in $G(x)$ for some x is a necessary condition for the existence of several fixed points in the dynamics (the sign of a circuit being defined as the product of the signs of its edges), and that the existence of a negative circuit is a necessary condition for the existence of an attractive cycle. These two results are inspired by rules for discrete models of genetic regulatory networks proposed by the biologist R. Thomas. The proof of the first result is modelled after a recent proof of the discrete Jacobian conjecture.

Key words: Discrete dynamical systems, Boolean networks, Regulatory networks, Genetic regulation, Differentiation, Homeostasis, Thomas' rules, Discrete Jacobian matrix, Jacobian conjecture
PACS: 37N25, 39A12, 92C40, 92D10

[1] Research partly supported by the French Ministry of Research (ACI IMPbio) and the Agence Nationale de la Recherche (ANR-05-JCJC-0126-01). A preliminary version of this article appears as preprint IML 2005-08 of the Institut de Mathématiques de Luminy.

1 Introduction

This article deals with properties of Boolean dynamical systems arising in biology.

Biologists often represent the results of their genetic and molecular investigations in terms of graphs. General information on molecular biology may be found in [2] or in the mathematical article [5]. In particular, genetic regulatory networks are usually represented by graphs, where vertices denote genes or regulatory products (e.g., RNA, proteins) whereas edges denote regulatory interactions between these genes or their products [16,6,25]. Regulatory interactions are further directed and signed ($+1$ or -1) to denote activatory versus inhibitory effects.

In order to relate regulatory networks to relevant dynamical properties, biologists often use them as a basis to generate dynamical models, using either a differential framework or a discrete framework [34,6]. For instance, in a differential model, the activity of a gene in a specific cell is measured by the concentration of the transcribed RNA in the cell, a quantity called the expression level of the gene, and the expression levels of n genes are modelled by an n-tuple $x \in \mathbb{R}^n$ obeying a differential equation $\dot{x} = f(x)$. As available data suggest that many interactions have to be modelled in terms of non-linear functions, typically with strong threshold effects, f is usually non-linear. It should be observed that the correspondence from regulatory networks to dynamics is not a function, and from the mathematical viewpoint, it is more satisfactory to turn the correspondence around and associate graphs to a given dynamics. Anyway, the biological pertinence of the model considered is evaluated by comparing numerical simulations with experimental observations, for instance biochemical characterizations of cellular states, phenotypes of genetic mutants, etc.

The biologist R. Thomas has proposed two rules relating the structure of regulatory networks to their dynamical properties [33]:

(1) a necessary condition for multistability (i.e., the existence of several stable fixed points in the dynamics) is the existence of a positive circuit in the regulatory network (the sign of a circuit being defined as the product of the signs of its edges);
(2) a necessary condition for the existence of an attractive cycle in the dynamics is the existence of a negative circuit.

These two types of dynamical properties correspond to important biological phenomena: cell differentiation processes in the first case, homeostasis or periodic behaviours (e.g., cell cycle or circadian rhythms) in the second case.

During the last decade, several authors have proposed demonstrations of these rules in the differential framework [18,29,11,30]. However, these demonstrations do not encompass the discrete framework, which initially nourished the intuition of R. Thomas. Discrete approaches such as Boolean networks are increasingly used in biology [13,32,34,14,35,27,1] because of the qualitative nature of most experimental data, together with a wide occurrence of non-linear regulatory relationships (e.g., combinatorial arrangements of molecular bindings, existence of cooperative or antagonist regulatory effects). Recently, [3] proved Thomas' rules in the monotonous Boolean case, i.e., when the dynamics of a system of n genes is given by a monotonous function $f : \{0,1\}^n \to \{0,1\}^n$ (see also [4]). Furthermore, an extensive analysis of the properties of discrete isolated circuits is provided in [19].

In this paper, we propose discrete counterparts of Thomas' rules in the general Boolean case and demonstrate them. In the differential framework, C. Soulé [30] associates to each state a signed directed graph, which is defined from the Jacobian matrix. From the biological viewpoint, this local character of regulatory networks is consistent with the fact that interactions are often context-sensitive, i.e., the effect of one regulatory product on a given gene depends on the presence of other regulatory products. We follow this approach in Section 2 by associating to a map $f : \{0,1\}^n \to \{0,1\}^n$ and a state $x \in \{0,1\}^n$ a signed directed graph $G(x)$, which is related to the discrete Jacobian matrix $J(x)$ defined in [22,23].

In Section 3, we state and prove Thomas' rule for the existence of positive circuits in the Boolean model (Theorem 3.2). Our proof of this rule is modelled after a recent proof by M.-H. Shih and J.-L. Dong of the discrete version of the Jacobian conjecture [26].

In Section 4, we state and prove a version of Thomas' rule relating the presence of attractive cycles in the dynamics to the existence of a negative circuit (Theorem 4.4), and we explore possible variants of these results. In particular, we show that the existence of some circuit, negative or not, follows from the presence of (non necessarily attractive) cycles (Theorem 4.5). It is worth observing that Theorems 3.2 and 4.4 give information on the vertices (i.e., the genes) involved in the observed dynamic behaviour.

As we believe that these results are of intrinsic interest, Sections 2, 3 and 4 are essentially written from the point of view of dynamical systems, almost without any reference to biology. The relationship with genetic regulatory networks is explained in Section 2.3.

In Section 5, Thomas' rules are illustrated with a simple example and further discussed in relation with the problem of regulatory network inference in molecular genetics and functional genomics.

Acknowledgments. We are very grateful to Christophe Soulé for stimulating discussions.

2 Signed directed graphs and Boolean dynamics

Let us start with preliminary notations. For $\beta \in \{0,1\}$, we define $\overline{\beta}$ by $\overline{0} = 1$ and $\overline{1} = 0$. Let n be a positive integer. For $x \in \{0,1\}^n$ and $I \subseteq \{1,\ldots,n\}$, $\overline{x}^I \in \{0,1\}^n$ is defined by:

$$(\overline{x}^I)_i = \begin{cases} x_i & \text{for } i \notin I, \\ \overline{x_i} & \text{for } i \in I. \end{cases}$$

When $I = \{i\}$ is a singleton, $\overline{x}^{\{i\}}$ is denoted by $\overline{x}^i$. The distance $d : \{0,1\}^n \times \{0,1\}^n \to \{0,1,\ldots,n\}$ is the Hamming distance: $d(x,y)$ is the number of $i \in \{1,\ldots,n\}$ such that $x_i \neq y_i$.

2.1 Boolean dynamics

We consider $\{0,1\}^n$ as the set of *states* of a dynamical system. In differential dynamical systems, the dynamics is typically governed by a differential equation $\dot{x} = f(x)$ where the map f is a vector field on, say, $\mathbb{R}^n$. In contrast, in discrete models, a dynamics is a binary relation R which we assume to be irreflexive: R gives the rule for updating a state, i.e., it is the set of pairs of states (x,y) such that state x can lead to state y. In particular, a *stable state* is a state x such that for no y, $(x,y) \in R$.

In the context considered in this paper (genetic networks), it is not realistic to assume a simultaneous update of all variables. Indeed, the Boolean dynamical systems we are interested in can be seen as discretizations of piecewise-linear differential systems [9,33,6,31], and for these systems, the set of trajectories meeting more than one threshold hyperplane at a time has measure 0. We shall therefore consider *asynchronous dynamics*, i.e., relations R such that:

$$(x,y) \in R \text{ implies } d(x,y) = 1,$$

i.e., $y = \overline{x}^i$ for some i. Such an asynchronous dynamics R may be non-deterministic (it needs not be a function), but even then, it is possible and convenient to represent it by a map

$$f : \{0,1\}^n \to \{0,1\}^n$$

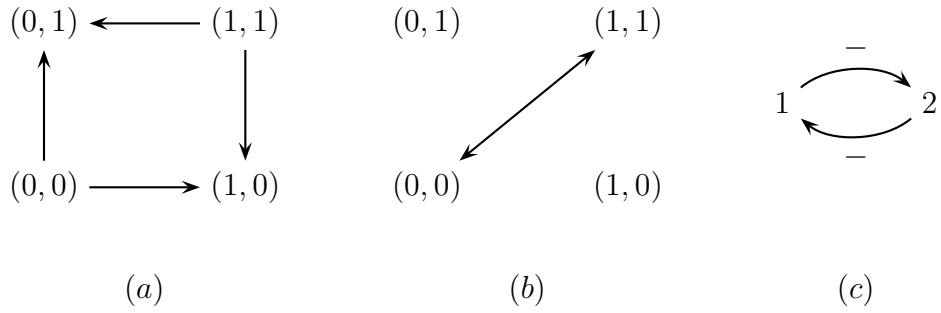


Fig. 1. (*a*) Asynchronous dynamics: the states of a system consisting in two variables 1 (horizontal axis) and 2 (vertical axis) are pictured; an arrow from state x to state $\overline{x}^i$ means that $f_i(x) \neq x_i$. (*b*) The associated map f. (*c*) The regulatory graph $G(x)$, which turns out not to depend on x.

with coordinate functions $f_1, \ldots, f_n$, defined by:

$$f_i(x) \neq x_i \text{ when } (x, \overline{x}^i) \in R.$$

Observe that a stable state is then a fixed point x for f ($f(x) = x$). Given such a map f, the corresponding asynchronous dynamics is defined in a straightforward way. The relation between an asynchronous dynamics and the corresponding map f is illustrated by a very simple example in Figure 1.

A *trajectory* in the dynamics is a sequence of states $(x^1, \ldots, x^r)$ such that for each $i = 1, \ldots, r-1$, $(x^i, x^{i+1}) \in R$. In terms of f, this means that for each i there exists $\varphi(i) \in \{1, \ldots, n\}$ such that

$$f_{\varphi(i)}(x^i) \neq x^i_{\varphi(i)} \text{ and } x^{i+1} = \overline{x^i}^{\varphi(i)}.$$

A trajectory $T = (x^1, \ldots, x^r)$ is completely described by its starting point x^1 and the map $\varphi : \{1, \ldots, r-1\} \to \{1, \ldots, n\}$, called its *strategy* [22,23], thus by abuse of notation, we shall write $T = (x^1, \varphi)$ as well.

A *cycle* is a trajectory of the form $(x^1, \ldots, x^r, x^1)$ with $r \geqslant 2$. A cycle C is completely described by one of its points, say x^1, and its strategy $\varphi : \{1, \ldots, r\} \to \{1, \ldots, n\}$, so we shall write $C = (x^1, \varphi)$ again. Observe that a trajectory (x^1, φ) with strategy φ is a cycle if, and only if, for any $i = 1, \ldots, n$, the cardinality of $\varphi^{-1}(i)$ is even (or zero); as a consequence, r is then even. We shall be especially interested in a specific class of cycles which correspond to periodic oscillations: a cycle (x^1, φ) is said to be *attractive* when no trajectory may leave it, i.e., for all $i = 1, \ldots, r$, $d(x^i, f(x^i)) = 1$. Equivalently, for all $i = 1, \ldots, r$,

$$f(x^i) = \overline{x^i}^{\varphi(i)}.$$

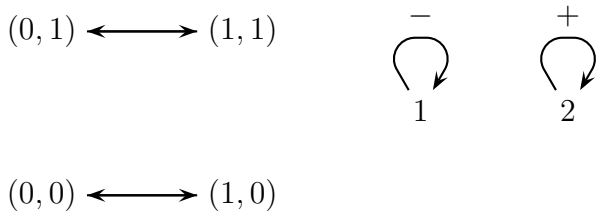


Fig. 2. A dynamics with no fixed point but a positive loop in the (constant) regulatory graph. The notation is the same as in Figure 1.

Figure 2 shows an example of dynamics with two attractive cycles:

$$((0,0),(1,0),(0,0)) \text{ and } ((0,1),(1,1),(0,1)),$$

Further examples of cycling dynamics are provided in Sections 4 and 5.

2.2 Discrete Jacobian matrices and signed directed graphs

Given $f : \{0,1\}^n \to \{0,1\}^n$, we attach to each $x \in \{0,1\}^n$ its *discrete Jacobian matrix* $J(x)$ as defined in [22,23]: $J(x)$ is the $n \times n$ matrix with (i,j)-entry

$$J(x)_{i,j} = \begin{cases} 1 & \text{if } f_i(\overline{x}^j) \neq f_i(x), \\ 0 & \text{otherwise.} \end{cases}$$

A *signed directed graph* is a directed graph with a sign, $+1$ or -1, attached to each edge.

Definition 2.1 *Given $f : \{0,1\}^n \to \{0,1\}^n$ and $x \in \{0,1\}^n$, define $G(x)$ to be the signed directed graph with vertex set $\{1,\ldots,n\}$ and with an edge from j to i when $J(x)_{i,j} = 1$, with positive sign when*

$$x_j = f_i(x),$$

and negative sign otherwise.

A *circuit* in a graph G is a non-empty sequence $(n_1,\ldots,n_k,n_1)$ of vertices such that G contains an edge from n_i to n_{i+1} for $i = 1,\ldots,k-1$ and an edge from n_k to n_1. If $I \subseteq \{1,\ldots,n\}$, an *$I$-circuit* is a circuit whose vertices belong to I. Observe that if $J \subseteq I$, a J-circuit is clearly an I-circuit. The *sign of a circuit C* is the product of the signs of its edges.

For instance, it is easy to check that the function f corresponding to the dynamics in Figure 1 is given by $f(x) = (\overline{x_2}, \overline{x_1})$; the Jacobian matrix associated to any state x is therefore given by:

$$J(x) = \begin{pmatrix} \overline{x_2} + \overline{x_2} & x_2 + \overline{x_2} \\ x_1 + \overline{x_1} & \overline{x_1} + \overline{x_1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where the sum here is the sum of $\{0, 1\}$ identified with the field $\mathbb{F}_2$. Therefore, the graph $G(x)$ at any state consists in a circuit between 1 and 2, hence a $\{1, 2\}$-circuit. Since $x_1 \neq f_2(x)$ and $x_2 \neq f_1(x)$, the two edges are negative and the circuit is positive.

2.3 Modelling genetic regulatory networks

In the context of genetic regulatory networks, we are interested in the evolution of the system consisting of n genes, which are denoted by the integers $1, \ldots, n$. Given a state $x = (x_1, \ldots, x_n) \in \{0, 1\}^n$, x_i denotes the (discretized) expression level of gene i. These expression levels are either 0 (when the gene product is considered absent or inactive) or 1 (when the gene product is present and active). Given a map $f : \{0, 1\}^n \to \{0, 1\}^n$, for each $x \in \{0, 1\}^n$ and $i = 1, \ldots, n$, $f_i(x)$ denotes the value to which x_i, the expression level of gene i, tends when the system is in state x.

The signed directed graph $G(x)$ attached to each state x encompasses a subset of the regulatory interactions found in the complete regulatory network.These graphs are analogous to the local interaction graphs considered in [30] for instance. Consequently, in our Boolean framework, a regulatory interaction and its sign may depend on the context, i.e., on the state of the system, in particular on the values of co-regulators acting on the same target. By taking the unions of graphs on states x, it is possible to lose some details and recover more global notions of regulatory networks as in [19,33]: for any $E \subseteq \{0, 1\}^n$, let $G(E) = \bigcup_{x \in E} G(x)$ be the graph with a positive (resp. negative) edge from j to i when there exists $x \in E$ such that $G(x)$ contains a positive (resp. negative) edge from j to i. Note that this opens the possibility to have both a positive and a negative edge connecting the same pair of vertices.

Some applications of this discrete model of genetic networks can be found in [34,15,8,24] and references therein. We now turn to the main results of this paper, and we shall comment again on their biological significance in Section 5.

3 Multistability and positive circuits

Suppose $0 \leqslant k \leqslant n$, and I is a k-element subset of $\{1, \ldots, n\}$. Then each $x \in \{0,1\}^n$ generates an affine k-dimensional subspace $x[\![I]\!]$ of $\{0,1\}^n = \mathbb{F}_2^n$ defined by:

$$x[\![I]\!] = \{y \in \{0,1\}^n \text{ such that } y_j = x_j \text{ for all } j \notin I\}.$$

We call such a subspace $x[\![I]\!]$ an *I-subcube*, or a *k-subcube*, or simply a *subcube* [26]. If κ is an I-subcube, a *κ-fixed point* is an $x \in \kappa$ such that $f_i(x) = x_i$ for all $i \in I$.

Lemma 3.1 *Let $f : \{0,1\}^n \to \{0,1\}^n$ and $I \subseteq \{1, \ldots, n\}$. If for each $x \in \{0,1\}^n$, $G(x)$ has no positive I-circuit, then for each I-subcube κ, f has at most one κ-fixed point.*

Proof — Proceed by induction on the cardinality k of I. The result holds trivially for $k = 0$, since a 0-subcube is a singleton. For $k = 1$, a 1-subcube is of the form $\kappa = \{x, \overline{x}^i\}$ for some x, i: if both x and $\overline{x}^i$ were κ-fixed points, then $f_i(x) = x_i \neq \overline{x_i} = f_i(\overline{x}^i)$, thus $G(x)$ would contain a positive loop on i, hence a positive $\{i\}$-circuit.

Now, if $1 \leqslant k \leqslant n-1$, let $\kappa = x[\![I]\!]$ be a $(k+1)$-subcube. Suppose for a contradiction that for each $x \in \{0,1\}^n$, $G(x)$ has no positive I-circuit, but that f has at least two κ-fixed points a and b. There are two cases:

- If $d(a,b) = r \leqslant k$, there is an r-element set J strictly included in I such that $b = \overline{a}^J$, therefore a and b both belong to the r-subcube $\lambda = a[\![J]\!]$ and are obviously λ-fixed points. Since for each $x \in \{0,1\}^n$, $G(x)$ has no positive I-circuit, $G(x)$ has no positive J-circuit, and we have a contradiction with the induction hypothesis.
- If $d(a,b) = k+1$, then $b = \overline{a}^I$. For each $i \in I$, $\overline{b}^i \in \kappa_i = a[\![I \setminus \{i\}]\!]$, a k-subcube. Now, a is a κ_i-fixed point for each $i \in I$, and $\overline{b}^i \neq a$ since $k+1 \geqslant 2$. For each $x \in \{0,1\}^n$, $G(x)$ has no positive $(I \setminus \{i\})$-circuit, thus the induction hypothesis implies that $\overline{b}^i$ is not a κ_i-fixed point for any $i \in I$. Hence there is a $j \in I \setminus \{i\}$ such that $f_j(\overline{b}^i) \neq (\overline{b}^i)_j$. Furthermore, $(\overline{b}^i)_j = b_j$ because $i \neq j$, and $b_j = f_j(b)$ because b is a κ-fixed point, so $f_j(\overline{b}^i) \neq f_j(b)$ and $J(b)_{j,i} = 1$. Therefore, the principal submatrix of $J(b)$ consisting of the (i,j)-entries for $i, j \in I$ has no zero column. As argued in [22,23], this implies that $G(b)$ has an I-circuit $C = (i_1, \ldots, i_\ell, i_1)$ with $i_1, \ldots, i_\ell \in I$ (column i_1 has a non-zero entry on some line i_2, and then going on with

column i_2, the sequence has to loop). Therefore, we have:

$$\begin{aligned} b_{i_2} &= f_{i_2}(b) = \sigma_1(b_{i_1}) \\ &\vdots \\ b_{i_\ell} &= f_{i_\ell}(b) = \sigma_{\ell-1}(b_{i_{\ell-1}}) \\ b_{i_1} &= f_{i_1}(b) = \sigma_\ell(b_{i_\ell}), \end{aligned}$$

where $\sigma_m : \{0,1\} \to \{0,1\}$ is either the map σ that exchanges 0 and 1 when the edge from i_m in C is negative, or the identity when the edge from i_m in C is positive. By hypothesis, the circuit C has to be negative, thus the number q of negative edges in C is odd, so $b_{i_1} = (\sigma_\ell \circ \cdots \circ \sigma_1)(b_{i_1}) = \sigma^q(b_{i_1}) = \overline{b_{i_1}}$, a contradiction.

□

As an immediate consequence, we have a general Boolean version of Thomas' rule relating multistability to positive circuits.

Theorem 3.2 *Let $f : \{0,1\}^n \to \{0,1\}^n$. If f has at least two fixed points, then there is an $x \in \{0,1\}^n$ such that $G(x)$ has a positive circuit. More precisely, if f has two fixed points a and b, and if I is such that $b = \overline{a}^I$, then there is an $x \in \{0,1\}^n$ such that $G(x)$ has a positive I-circuit.*

Proof — The first assertion follows from the second one. For a proof of the second assertion, it suffices to observe that a and b are two $a[\![I]\!]$-fixed points and to use Lemma 3.1. □

The requirement in Theorem 3.2 is clearly not a sufficient condition for multistability. For instance, the dynamics given in Figure 2 for $n = 2$ has no fixed point, whereas the regulatory graph associated to any state has a positive loop.

Recall that a graph is said to be *simple* when for any two vertices i and j, there is at most one edge from i to j. For each $x \in \{0,1\}^n$, define $\Gamma(x)$ to be the simple directed graph with vertex set $\{1,\ldots,n\}$ and an edge from j to i when $J(x)_{i,j} = 1$, i.e., the graph whose adjacency matrix if the transpose of $J(x)$. Observe that $\Gamma(x)$ is the graph underlying the signed directed graph $G(x)$. In [26], M.-H. Shih and J.-L. Dong prove a discretised version of the Jacobian conjecture, which relates the dynamical behaviour of f with the (non-signed) graphs $\Gamma(x)$. They show that if $f : \{0,1\}^n \to \{0,1\}^n$ is such that for each $x \in \{0,1\}^n$, $\Gamma(x)$ has no circuit, then f has a unique fixed point. The proof of Lemma 3.1 is modelled after the proof of Shih-Dong's theorem, and the contraposition of Theorem 3.2 gives a positive version of this result, with a

stronger hypothesis (no positive circuit) and a stronger conclusion (at most one fixed point).

4 Cycles and negative circuits

Thomas' rule relating homeostasis and negative circuits has to be reformulated in the discrete framework. This can be done in various ways. Should the cycle be attractive? Is it then possible to infer the existence of a negative circuit in the graph $G(x)$ attached to a single state x or to a set of states?

We prove in Section 4.2 that the existence of an attractive cycle C implies the presence of a negative circuit in the union of the graphs $G(x)$ for x a state of C (Theorem 4.4), and we observe that the attractiveness hypothesis is necessary. However, the existence of an arbitrary (non-necessarily attractive) cycle implies the existence of some (positive or negative) circuit (Theorem 4.5). By exhaustive exploration, one can further show that an attractive cycle implies the presence of a negative circuit in the graph $G(x)$ attached to a single state x when $n \leqslant 3$, but we have no proof that this holds for any n.

4.1 First recurrence function

If $r \in \mathbb{N}$, $r \geqslant 1$ and $k, \ell \in \{1, \ldots, r\}$ are such that $k \neq \ell$, $]k, \ell[$ denotes either $\{k+1, \ldots, \ell-1\}$ if $k < \ell$, or $\{k+1, \ldots, r, 1, 2, \ldots, \ell-1\}$ if $\ell < k$, and $[k, \ell[= \{k\} \cup]k, \ell[$. The intuition is that we view $1, \ldots, r$ as r points of a cycle.

Definition 4.1 *If X is a set and $\varphi : \{1, \ldots, r\} \to X$, the* first recurrence function *for φ is the partial function $R_\varphi : \{1, \ldots, r\} \to \{1, \ldots, r\}$ such that $R_\varphi(k)$ is the unique $\ell \neq k$ satisfying $\varphi(k) = \varphi(\ell)$ and for each $i \in]k, \ell[$, $\varphi(i) \neq \varphi(k)$, if such an ℓ exists.*

Clearly, $R_\varphi(k)$ is defined if, and only if, $\varphi(k)$ has at least two preimages under φ. In particular, if for all $x \in X$, $\varphi^{-1}(x)$ has cardinality at least 2, R_φ is a permutation of $\{1, \ldots, r\}$.

Lemma 4.2 *If $\varphi : \{1, \ldots, r\} \to X$ is such that for all $x \in X$, $\varphi^{-1}(x)$ has cardinality at least 2, then there exists $k \in \{1, \ldots, r\}$ such that $\{\varphi(k), \ldots, \varphi(R_\varphi(k) - 1)\} = \{\varphi(1), \ldots, \varphi(r)\} = X$.*

Proof — Otherwise, for each $k \in \{1, \ldots, r\}$, there is $i \in X$ such that

$$[k, R_\varphi(k)[\cap \varphi^{-1}(i) = \varnothing. \tag{1}$$

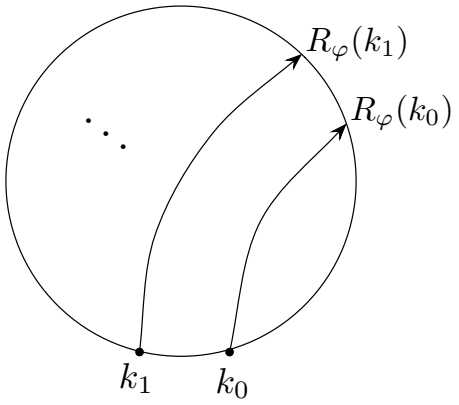


Fig. 3. First recurrence function in the proof of Lemma 4.2.

Take $k' \in \varphi^{-1}(i)$ such that:

$$]k', k[\cap \varphi^{-1}(i) = \varnothing. \tag{2}$$

Then $k' \neq k$ because $k \notin \varphi^{-1}(i)$. Moreover, by definition of R_φ, we have $R_\varphi(k') \in \varphi^{-1}(i)$. Therefore (1) implies $R_\varphi(k') \notin [k, R_\varphi(k)[$ and (2) implies $R_\varphi(k') \notin [k', k[$. Hence $[k, R_\varphi(k)[$ is a strict subset of $[k', R_\varphi(k')[$. By iterating this process, we may construct an infinite sequence $k_0, k_1, k_2, \ldots$ such that $k_i \neq k_{i+1}$ and $[k_i, R_\varphi(k_i)[$ is a strict subset of $[k_{i+1}, R_\varphi(k_{i+1})[$ for each $i \geqslant 0$: see Figure 3. But then, $(k_i)_{i \geqslant 0}$ is an infinite sequence of distinct elements of $\{1, \ldots, r\}$ and we have a contradiction. □

4.2 Attractive cycles

Lemma 4.3 *Assume $f : \{0,1\}^n \to \{0,1\}^n$ has an attractive cycle $(x^1, \ldots, x^r, x^1)$, with strategy $\varphi : \{1, \ldots, r\} \to \{1, \ldots, n\}$. Then R_φ is a permutation, and for each $k \in \{1, \ldots, r\}$ and each $i \in [k, R_\varphi(k)[$, $G(x^i)$ has an edge from $\varphi(i)$ to $\varphi(i+1)$ with sign ε_i such that $\prod_{i \in [k, R_\varphi(k)[} \varepsilon_i = -1$.*

Proof — Under the conditions of the lemma, it is clear that the preimage of each element in the image of φ has an even non-zero cardinality. Therefore, R_φ is a permutation. Let then $k \in \{1, \ldots, r\}$ and $\ell = R_\varphi(k)$, and let $i \in [k, \ell[$. We take indices modulo r, i.e., we identify $r + j$ and j. Since the cycle $(x^1, \ldots, x^r, x^1)$ is attractive, $f(x^i) = x^{i+1}$ and $f(x^{i+1}) = x^{i+2}$, hence:

$$f_{\varphi(i+1)}(x^i) = x^{i+1}_{\varphi(i+1)} \tag{3}$$

and

$$f_{\varphi(i+1)}(x^{i+1}) = x^{i+2}_{\varphi(i+1)}. \tag{4}$$

By definition of the strategy φ, we have:

$$x^{i+1} = \overline{x^i}^{\varphi(i)} \tag{5}$$

and

$$x^{i+2} = \overline{x^{i+1}}^{\varphi(i+1)}. \tag{6}$$

By (5) and (4) we have:

$$f_{\varphi(i+1)}\left(\overline{x^i}^{\varphi(i)}\right) = f_{\varphi(i+1)}(x^{i+1}) = x^{i+2}_{\varphi(i+1)},$$

and by (6) and (3) we have:

$$x^{i+2}_{\varphi(i+1)} \neq x^{i+1}_{\varphi(i+1)} = f_{\varphi(i+1)}(x^i).$$

As a consequence:

$$f_{\varphi(i+1)}\left(\overline{x^i}^{\varphi(i)}\right) \neq f_{\varphi(i+1)}(x^i)$$

and $G(x^i)$ has an edge from $\varphi(i)$ to $\varphi(i+1)$. Let ε_i be its sign: $\varepsilon_i = -1$ if, and only if, $x^i_{\varphi(i)} \neq f_{\varphi(i+1)}(x^i) = x^{i+1}_{\varphi(i+1)}$. Hence:

$$x^{i+1}_{\varphi(i+1)} = \left(x^i_{\varphi(i)}\right)^{\varepsilon_i},$$

where, for $\beta \in \{0,1\}$, $\beta^{+1} = \beta$ and $\beta^{-1} = \overline{\beta}$. This holds for any $i \in [k, \ell[$, therefore:

$$x^{\ell}_{\varphi(\ell)} = \left(x^k_{\varphi(k)}\right)^{\prod_{i\in[k,\ell[}\varepsilon_i}.$$

Since in addition $\ell = R_\varphi(k)$ we have $x^\ell_{\varphi(\ell)} = x^\ell_{\varphi(k)} \neq x^k_{\varphi(k)}$, and consequently $\prod_{i\in[k,\ell[}\varepsilon_i = -1$. □

Theorem 4.4 *If $f : \{0,1\}^n \to \{0,1\}^n$ has an attractive cycle $C = (x^1, \ldots, x^r, x^1)$ with strategy φ, then $G(C) = G(x^1) \cup \cdots \cup G(x^r)$ has a negative circuit with vertices $\varphi(1), \ldots, \varphi(r)$.*

Proof— Since $(x^1, \ldots, x^r, x^1)$ is a cycle, φ enjoys the condition of Lemma 4.2, thus there exists k such that $\{\varphi(k), \ldots, \varphi(R_\varphi(k)-1)\} = \{\varphi(1), \ldots, \varphi(r)\}$. By Lemma 4.3 and the remark that $\varphi(R_\varphi(k)) = \varphi(k)$, $C = (\varphi(k), \ldots, \varphi(R_\varphi(k) - 1), \varphi(R_\varphi(k)))$ is a negative circuit of $G(C) = \bigcup_{i=1}^r G(x^i)$, and its vertex set is $\{\varphi(1), \ldots, \varphi(r)\}$. □

One could think of other notions of cycles than the cycles of the asynchronous dynamics considered here. For instance, we could consider a form of synchronous dynamics in which all variables are updated simultaneously, and define an *f-cycle* in the obvious way, i.e., to be a sequence $(x^1, \ldots, x^r, x^1)$ of states such that $f(x^i) = x^{i+1}$ for $i < r$ and $f(x^r) = x^1$. For example, the

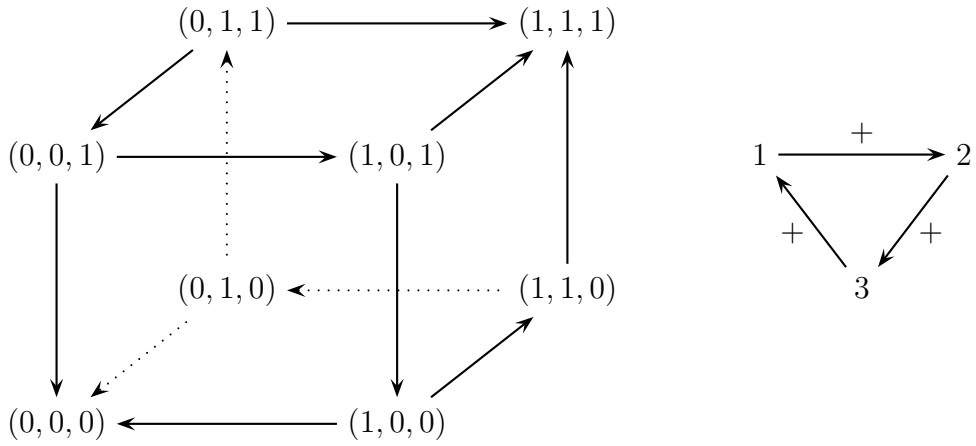


Fig. 4. A dynamics with no attractive cycle, a non-attractive one, and no negative circuit in the (constant) regulatory graph. The notation is the same as in Figure 1; dotted lines are used to ease the 3-cube visualisation.

synchronous dynamics for the cross-inhibitory circuit described in Figure 1 has an f-cycle $((0,0),(1,1),(0,0))$, while the graph $G(x)$ associated to any state x has no negative circuit. Therefore, the synchronous dynamics does not comply with Theorem 4.4. In contrast, the asynchronous dynamics associated with the same graph has no cycle, which is consistent with Theorem 4.4.

Building on these results, generalizations of Theorems 3.2 and 4.4 can be found in [20,21].

4.3 Arbitrary cycles

The dynamics given in Figure 4 with $n = 3$ is characterised by the unique, non-attractive cycle

$$((1,0,0),(1,1,0),(0,1,0),(0,1,1),(0,0,1),(1,0,1),(1,0,0))$$

and by two fixed points $(0,0,0)$ and $(1,1,1)$. This dynamics is essentially taken from [19]. The regulatory graph associated to any state is a single positive circuit, thus it has no negative circuit. The attractiveness hypothesis is therefore necessary to conclude with the presence of a negative circuit. Yet, the following theorem asserts that the presence of some circuit, negative or not, follows from the existence of (non necessarily attractive) cycles.

Theorem 4.5 *If $f : \{0,1\}^n \to \{0,1\}^n$ has a cycle $C = (x^1, \ldots, x^r, x^1)$ with strategy φ, then $G(C) = G(x^1) \cup \cdots \cup G(x^r)$ has a circuit.*

Proof — Suppose that $C = (x^1, \ldots, x^r, x^1) = (x^1, \varphi)$ is a cycle of minimal length, i.e., such that there is no cycle of length strictly smaller than r. Let $I \subseteq \{1, \ldots, n\}$ be the image of φ. In order to conclude that $G(C)$ has a circuit, it suffices to prove that for any $j \in I$, there exist $i \in I$ and $x \in \{x^1, \ldots, x^r\}$ such that $J(x)_{i,j} = 1$.

Fix $j \in I$ and suppose for a contradiction that for all $i \in I$ and for all x, $f_i(x) = f_i(\overline{x}^j)$. Let k be such that $\varphi(k) = j$, and let $\ell = R_\varphi(k)$. For $p = k, \ldots, \ell - 1$, define y^p by $y^k = x^k$ and for $p = k, \ldots, \ell - 2$:

$$y^{p+1} = \overline{y^p}^{\varphi(p+1)}.$$

Let us show that $(y^k, \ldots, y^{\ell-1})$ is a trajectory. By definition of the first recurrence function R_φ, for $p = k, \ldots, \ell - 1$, we have

$$x^p = \overline{x^k}^{\{\varphi(k),\ldots,\varphi(p-1)\}} \text{ and } y^p = \overline{y^k}^{\{\varphi(k+1),\ldots,\varphi(p)\}},$$

therefore:

$$y^p = \overline{x^{p+1}}^{\varphi(k)} = \overline{x^{p+1}}^j. \tag{7}$$

Now, $(x^{k+1}, \ldots, x^\ell)$ is a trajectory, thus for $p = k, \ldots, \ell - 2$, we have

$$f_{\varphi(p+1)}(x^{p+1}) \neq x^{p+1}_{\varphi(p+1)}.$$

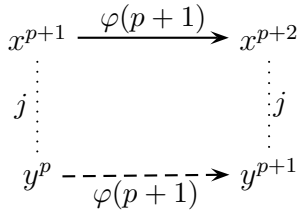
By the hypothesis on j, this implies that:

$$f_{\varphi(p+1)}(y^p) \neq x^{p+1}_{\varphi(p+1)},$$

and $y^p_{\varphi(p+1)} = \left(\overline{x^{p+1}}^j\right)_{\varphi(p+1)} = x^{p+1}_{\varphi(p+1)}$ since $\ell = R_\varphi(k)$. Therefore:

$$f_{\varphi(p+1)}(y^p) \neq y^p_{\varphi(p+1)},$$

and (y^p, y^{p+1}) is a trajectory for $p = k, \ldots, \ell-2$, as illustrated in the following diagram.



By composing these trajectories of length 1, we obtain the expected trajectory $(y^k, \ldots, y^{\ell-1})$. Now, $y^k = x^k$ and by (7) we have:

$$y^{\ell-1} = \overline{x^\ell}^{\varphi(k)} = \overline{x^\ell}^{\varphi(\ell)} = x^{\ell+1}.$$

Hence we have constructed a cycle

$$(x^1, \ldots, x^k = y^k, y^{k+1}, \ldots, y^{\ell-1} = x^{\ell+1}, \ldots, x^r, x^1)$$

of length $r - 2 < r$, in contradiction with the minimality hypothesis. □

5 Illustration and discussion

In this section, we illustrate the notions presented in the previous sections through a simple example for which we give detailed computations, and we discuss these notions in relation with the problem of regulatory network inference.

Figure 5 describes the Boolean asynchronous dynamics of a simple three-genes genetic regulatory network. This dynamics has a single fixed point $(0, 0, 1)$ and an attractive cycle

$$((1,0,0),(1,0,1),(1,1,1),(1,1,0),(1,0,0)) = (x^1, x^2, x^3, x^4, x^1).$$

The strategy of this cycle is the map $\varphi : \{1,2,3,4\} \to \{1,2,3\}$ which associates to i the variable $\varphi(i)$ updated at state x^i. Since

$$x^2 = \overline{x^1}^3, x^3 = \overline{x^2}^2, x^4 = \overline{x^3}^3, x^1 = \overline{x^4}^2,$$

we have $\varphi : 1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 3, 4 \mapsto 2$. Let us compute the discrete Jacobian matrix $J(x)$ of the map f associated to this dynamics. It is not difficult to check that:

$$f(x) = (x_1, x_1x_3, \overline{x_2}).$$

Therefore:

$$J(x) = \begin{pmatrix} x_1 + \overline{x_1} & 0 & 0 \\ x_1x_3 + \overline{x_1}x_3 & 0 & x_1x_3 + x_1\overline{x_3} \\ 0 & \overline{x_2} + x_2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ x_3 & 0 & x_1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{8}$$

In particular,

$$J(1,0,0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, J(0,0,1) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } J(1,1,1) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

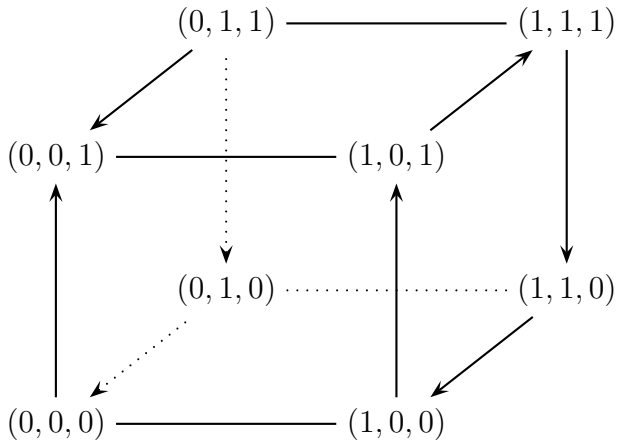
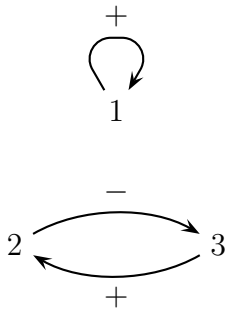


Fig. 5. The asynchronous dynamics of a three-element genetic regulatory network. The notation is the same as in Figure 4.

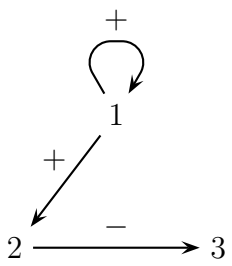
The local regulatory network $G(1,0,0)$ thus contains three edges, one from 2 to 3, one from 3 to 2 and a loop on 1. For $x = (1,0,0)$, $x_1 = f_1(x)$, $x_2 = 0 \neq f_3(x)$ and $x_3 = 0 = f_2(x)$, thus the loop on gene 1 is positive (self-activation), gene 2 inhibits gene 3 (negative edge), and gene 3 activates gene 2 (positive edge). Therefore:

$G(1,0,0)$ =



Similarly, the local regulatory network $G(0,0,1)$ contains three edges, one from 1 to 2, one from 2 to 3 and a loop on 1. For $x = (0,0,1)$, $x_1 = f_1(x) = 0 = f_2(x)$ and $x_2 = 0 \neq f_3(x)$; consequently, the edge from 2 to 3 is negative, whereas the two other edges are positive.

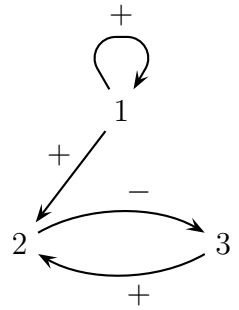
$G(0,0,1)$ =



According to (8), the regulatory network $G(1,1,1)$ is simply the union of the

graphs just mentioned: $G(1,1,1) = \bigcup_{x\in\{0,1\}^3} G(x) = G(1,0,0) \cup G(0,0,1)$.

$G(1,1,1)$ =



In agreement with Theorem 4.4, the regulatory graph $G(1,0,0) \cup G(1,0,1) \cup G(1,1,1) \cup G(1,1,0) = G(1,1,1)$ contains a negative circuit, and the strategy φ of the attractive cycle corresponding to the oscillatory behaviour of 2 and 3 enables us to infer that the vertices of this circuit are 2 and 3.

The regulatory network $G(x)$ associated to each state x has, by (8) and the equality $x_1 = f_1(x)$, a positive self-regulation of 1. The dynamics illustrated in Figure 5 does not exhibit multiple fixed points, but rather two alternative attractors: the attractive cycle and the fixed point $(0,0,1)$. Thus, in this example, we do not have an exact application of Theorem 3.2 on the requirement of a positive circuit in presence of multistability, but rather an illustration of a generalization of this theorem to the coexistence of alternative attractors, which is achieved in [20].

From a biological point of view, it is interesting to note that similar negative circuits are found at the core of cell cycle and circardian rhythm controlling networks: see, e.g., [10,12] and references therein. Furthermore, in our example, the dynamical role of the negative circuit depends on the presence of some regulatory product (the product of gene 1), thereby defining what we may call a *functionality context* (the set of states x for which $G(x)$ contains the circuit), which corresponds to the biological notion of check point. In the example, the functionality context of the negative circuit is given by $x_1 = 1$.

For dimensions higher than 3, the computation of the regulatory network requires the computation of $G(x)$ for larger numbers of states x in order to cover the space $\{0,1\}^n$. However, for higher dimensions, it should still be possible to analyse the discrete Jacobian matrix around specific sets of states and induce the corresponding local regulatory networks, leading to the progressive delineation of the feedback circuits present in the original regulatory network. This corresponds to the problem of the inference of genetic regulatory networks from temporal gene expression data obtained at the level of transcription (e.g., using DNA chips) or proteins (approaches combining protein chromatographic separation and mass spectrometry). An interesting strategy for the inference of genetic regulatory networks from kinetic data is proposed by [17]. In this context, the main challenge consists in dealing with limited and imperfect dy-

namical data sets, i.e., sparse, poorly reproducible, or imprecise time series, for relatively small subsets of possible initial conditions. In this respect, it should be observed that our tentative approach focuses on structural properties of the inferred network (occurrence of signed circuits, list of genes involved in these circuits), which could still largely hold even though some intermediary elements acting in the original networks could be missed.

Finally, we have to concede that, at this point, we have not considered explicitly any kind of time constraints on concurrent transitions under the asynchronous updating assumption. Such time constraints can lead to the elimination of some edges from the full asynchronous state transition graph. For instance, [7,28] consider different temporization approaches applied to biological systems. Consequently, the number and the type of attractors corresponding to a given regulatory network can be affected. However, once a given dynamical property is observed, our procedure for the inference of the underlying graph remains relevant.

References

[1] R. Albert, H. Othmer, The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *drosophila melanogaster*, J. Theoret. Biol. 223 (2002) 1–18.

[2] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J. D. Watson, Molecular biology of the cell, Garland Publishing, 1994.

[3] J. Aracena, Modèles mathématiques discrets associés à des systèmes biologiques. application aux réseaux de régulation génétiques, Thèse de doctorat, Université Joseph Fourier, Grenoble (2001).

[4] J. Aracena, J. Demongeot, E. Goles, On limit cycles of monotone functions with symmetric connection graph, Theoret. Comput. Sci. 322 (2) (2004) 237–244.

[5] A. Carbone, M. Gromov, Mathematical slices of molecular biology, Gaz. Math. 88 (2001) 11–80.

[6] H. de Jong, Modeling and simulation of genetic regulatory systems: A literature review, J. Comput. Biol. 9 (1) (2002) 67–103.

[7] A. Fauré, A. Naldi, C. Chaouiya, D. Thieffry, Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle, Bioinformatics 22 (14) (2006) 124–131.

[8] A. Ghysen, R. Thomas, The formation of sense organs in Drosophila: a logical approach, BioEssays 25 (2003) 802–807.

[9] L. Glass, S. A. Kauffman, The logical analysis of continuous non-linear biochemical control networks, J. Theoret. Biol. 39 (1973) 103–129.

[10] A. Golbeter, Computational approaches to cellular rhythms, Nature 420 (2002) 238–245.

[11] J.-L. Gouzé, Positive and negative circuits in dynamical systems, J. Biol. Systems 6 (1998) 11–15.

[12] A. C.-N. J. J. Tyson, B. Novak, The dynamics of cell cycle regulation, BioEssays 24 (2002) 1095–1109.

[13] S. A. Kauffman, Metabolic stability and epigenesis in randomly constructed genetic nets, J. Theoret. Biol. 22 (1969) 437–467.

[14] S. A. Kauffman, The origins of order: Self-organization and selection in evolution, Oxford University Press, 1993.

[15] L. Mendoza, D. Thieffry, E. R. Alvarez-Buylla, Genetic control of flower morphogenesis in arabidopsis thaliana: a logical analysis, Bioinformatics 15 (1999) 593–606.

[16] J. Monod, F. Jacob, General conclusions: Teleonomic mechanisms in cellular metabolism, growth and differentiation, in: Cold Spring Harbor Symp. Quantitative Biol., vol. 26, 1961.

[17] T. J. Perkins, M. Hallett, L. Glass, Inferring models of gene expression dynamics, J. Theoret. Biol. 230 (3) (2004) 289–299.

[18] E. Plahte, T. Mestl, S. W. Omholt, Feedback loops, stability and multistationarity in dynamical systems, J. Biol. Systems 3 (1995) 409–413.

[19] É. Remy, B. Mossé, C. Chaouiya, D. Thieffry, A description of dynamical graphs associated to elementary regulatory circuits, Bioinformatics 19 (2) (2003) 172–178.

[20] É. Remy, P. Ruet, On differentiation and homeostatic behaviours of Boolean dynamical systems, in: Trans. Comput. Systems Biol., vol. 4780 of Lect. Notes Comput. Sci., Springer, 2007.

[21] É. Remy, P. Ruet, D. Thieffry, Positive or negative regulatory circuit inference from multilevel dynamics, in: Positive Systems: Theory and Applications, vol. 341 of Lect. Notes Control Inform. Sci., Springer, 2006.

[22] F. Robert, Discrete iterations: a metric study, vol. 6 of Ser. Comput. Math., Springer, 1986.

[23] F. Robert, Les systèmes dynamiques discrets, vol. 19 of Mathématiques et Applications, Springer, 1995.

[24] L. Sánchez, D. Thieffry, Segmenting the fly embryo: a logical analysis of the pair-rule cross-regulatory module, J. Theoret. Biol. 224 (2003) 517–537.

[25] T. Schlitt, A. Brazma, Current approaches to gene regulatory network modelling, BMC Bioinformatics 8 (Suppl. 6) (2007) S9.

[26] M.-H. Shih, J.-L. Dong, A combinatorial analogue of the Jacobian problem in automata networks, Adv. in Appl. Math. 34 (1) (2005) 30–46.

[27] I. Shmulevich, E. R. Dougherty, W. Zhang, From Boolean to probabilistic Boolean networks as models of genetic regulatory networks, Proc. IEEE 90 (11) (2002) 1778–1792.

[28] H. Siebert, A. Bockmayr, Incorporating time delays into the logical analysis of gene regulatory networks, in: Comput. Methods Systems Biol., vol. 4210 of Lect. Notes Comput. Sci., Springer, 2006.

[29] E. H. Snoussi, Necessary conditions for multistationarity and stable periodicity, J. Biol. Systems 6 (1998) 3–9.

[30] C. Soulé, Graphic requirements for multistationarity, ComPlexUs 1 (2003) 123–133.

[31] C. Soulé, Mathematical approaches to gene regulation and differentiation, C. R. Acad. Sci. Paris : Biol. 329 (1) (2006) 13–20.

[32] R. Thomas, Boolean formalization of genetic control circuits, J. Theoret. Biol. 42 (1973) 563–585.

[33] R. Thomas, On the relation between the logical structure of systems and their ability to generate multiple steady states and sustained oscillations, in: Ser. Synergetics, vol. 9, Springer, 1981, pp. 180–193.

[34] R. Thomas, R. D'Ari, Biological feedbacks, CRC Press, 1990.

[35] R. Thomas, D. Thieffry, M. Kaufman, Dynamical behaviour of biological regulatory networks I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state, Bull. Math. Biol. 57 (1995) 247–276.