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, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9 (France). Email: {remy,ruet}@iml.univ-mrs.fr

*Institut de Biologie du Développement de Marseille, CNRS - INSERM - Université de la Méditerranée, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9 (France). Email: thieffry@ibdm.univ-mrs.fr

To each Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}^n$ and each $x \in \{0,1\}^n$, we associate a signed directed graph $G(x)$, and we prove 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 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 rules have been proposed by the biologist R. Thomas in the course of his analysis of genetic regulatory networks. The proof of the first one relies on a recent proof of the discrete Jacobian conjecture.

Key words: Discrete dynamical systems, Regulatory networks, Genetic regulation, Differentiation, Homeostasis, Discrete Jacobian matrix, Jacobian conjecture

PACS: 37N25, 39A12, 92C40, 92D10

1 Introduction

This article deals with conjectures on dynamical systems raised in the context of biology.

1 Research partly supported by the ACI IMPbio Program of the French Ministry of Research.

Preprint submitted to Elsevier

25 April 2005
Biologists\(^2\) often represent the results of their genetic and molecular investigations into biological processes in terms of graphs \([5]\). For example, genetic regulatory networks are graphs where vertices represent genes or regulatory products (e.g., RNA, proteins) and edges represent interactions between these genes or their products. Regulatory interactions are further directed and signed (+1 or -1): the sign denotes inhibitory versus activatory effect.

In order to relate these regulatory graphs to the dynamical properties of a biological system, biologists often use them as a basis to generate dynamical models, using either a differential or a discrete framework. 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 a differential model for instance, \(x \in \mathbb{R}^n\) and obeys a differential equation \(\dot{x} = f(x)\). Available data suggest that the regulatory phenomenon is strongly non-linear, with threshold effects, hence \(f\) is usually a non-linear function, whose parameters need to be adjusted. Observe that the correspondence from graphs to dynamics is not a function, and from the mathematical viewpoint, it is more satisfactory to turn things around and build regulatory graphs from dynamics. Anyway, the biological pertinence of the model considered is then evaluated by comparing numerical simulations with experimental observations, for instance biochemical characterisation of cellular states, phenotypes of genetic mutants, etc.

The biologist R. Thomas has enunced two beautiful rules relating the structure of regulatory graphs to their dynamical properties \([16]\):

(i) 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 graph (the sign of a circuit being the product of the signs of its edges);
(ii) 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 of sustained oscillatory 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 \([8,14,6,15]\). However, these demonstrations do not encompass the discrete framework, which initially nourished the intuition of R. Thomas. Note that such discrete approaches are increasingly used in biology because of the qualitative nature of most experimental data.

\(^2\) For more information on molecular biology, we refer the reader to a standard textbook like \([1]\) or to the mathematical article \([3]\).
together with a wide occurrence of non-linear regulatory relationships (e.g., combinatorial arrangements of molecular bindings, existence of cooperative or antagonist regulatory effects). Two exceptions are a proof Thomas’ rules in the monotonous Boolean case, i.e., when the dynamics of a system of $n$ genes is given by a function $f : \{0, 1\}^n \to \{0, 1\}^n$ that is monotonous [2], and an exhaustive study of the discrete case when the regulatory graph is an isolated circuit [9].

In this paper, we provide discrete counterparts of Thomas’ rules in the general Boolean case and demonstrate them. In the differential framework, C. Soulé [15] associates to each state a regulatory graph, which is defined from the Jacobian matrix at that state. From the biological viewpoint, this local character of regulatory graphs 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 regulatory graph $G(x)$, which is related to the discrete Jacobian matrix $J(x)$ defined by Robert in [10,11]. In Section 3, we state and prove Thomas’ rule for the existence of positive circuits in the Boolean model (Theorem 3.2), by using a recent proof by M.-H. Shih and J.-L. Dong of the discrete version of the Jacobian conjecture [12]. In Section 4, we state and prove a version of Thomas’ rule for the existence of negative circuits (Theorem 4.4), and we study possible variants of these results. 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. These results are illustrated with a simple example in Section 5 and further discussed in relation with the problem of regulatory network inference in molecular genetics and functional genomics.

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

2 Regulatory graphs and dynamics

We introduce here the definitions which enable to associate regulatory graphs to a dynamics.

Let us start with preliminary notations. For $\beta \in \{0, 1\}$, we define $\overline{\beta}$ by $\overline{1} = 1$ and $\overline{0} = 0$. For $x \in \{0, 1\}^n$ and $I \subseteq \{1, \ldots, n\}$, $\overline{x}_I \in \{0, 1\}^n$ is defined by $(\overline{x}_I)_i = x_i$ for $i \notin I$ and $(\overline{x}_I)_i = \overline{x}_i$ for $i \in I$. When $I = \{i\}$ is a singleton, $\overline{x}_i$ is denoted by $\overline{x}_i$. The distance $d_1$ on $\{0, 1\}^n$ is the Hamming distance: $d_1(x, y)$ is the number of $i \in \{1, \ldots, n\}$ such that $x_i \neq y_i$.

Let $n$ be a positive integer. Genes are denoted by $g_1, \ldots, g_n$. A state of the
system is an $x = (x_1, \ldots, x_n) \in \{0, 1\}^n$, where $x_i$ is the (discretised) expression level of gene $g_i$. These expression levels are assumed to be either 0 (gene not expressed) or 1 (gene expressed). We shall be interested in the evolution of the system consisting in the $n$ genes. Consider a map $f : \{0, 1\}^n \to \{0, 1\}^n$, $f(x) = (f_1(x), \ldots, f_n(x))$. 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 $g_i$, tends when the system is in state $x$.

### 2.1 Dynamics

In a differential model of genetic regulation, for example with differential equation $\dot{x} = f(x)$, the dynamics is simply given by the function $f$. In contrast, in discrete models, a dynamics is a binary relation which, given $f : \{0, 1\}^n \to \{0, 1\}^n$, needs to be chosen. Given such an $f$ and a state $x$, we denote by $\text{Com}(x)$ the set of genes $\{g_1, \ldots, g_n\}$ which may switch (or “commute”) their expression level, i.e., those for which $x_i \neq f_i(x)$. Clearly, $\text{Com}(x) = \emptyset$ means that $x$ is a fixed point. Several rules for updating the state are possible, and two of them are more frequently used:

the *synchronous* dynamics: $\{(x, f(x)) \text{ s.t. } x \in \{0, 1\}^n\}$, where all the expression levels $x_i$ are simultaneously updated to $f_i(x)$ in one step;

the *fully asynchronous* dynamics: $\{(x, \overline{x}) \text{ s.t. } x \in \{0, 1\}^n, i \in \text{Com}(x)\}$, where the expression level of only one gene is updated at each step.

These two rules give rise to quite different dynamical behaviours, as illustrated in Figure 1 for the following very simple example:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$f(x)$</th>
<th>$\text{Com}(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0)</td>
<td>(1, 1)</td>
<td>${1, 2}$</td>
</tr>
<tr>
<td>(0, 1)</td>
<td>(0, 1)</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>(1, 0)</td>
<td>(1, 0)</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>(1, 1)</td>
<td>(0, 0)</td>
<td>${1, 2}$</td>
</tr>
</tbody>
</table>

In the Boolean model considered here, it is not reasonable to authorise several simultaneous updates: the asynchronous dynamics is more biologically realistic (see for instance the discretisation of piecewise-linear differential equations [16]), and more relevant from our viewpoint, as we shall see in Section 4. In the sequel, unless otherwise specified, the chosen dynamics is the asynchronous one.
Fig. 1. (a) Asynchronous dynamics: the states of a system consisting in two genes $g_1$ (horizontal axis) and $g_2$ (vertical axis) are pictured; an arrow from state $x$ to state $y$ means that $i \in \text{Com}(x)$. (a') Synchronous dynamics of the same system, i.e., the associated function $f$. (b) The regulatory graph $G(x)$, which here does not depend on $x$.

A path (for $f$) is a sequence of states $(x^1, \ldots, x^r)$ such that for each $i = 1, \ldots, r$, there exists $\varphi(i) \in \{1, \ldots, n\}$ such that

$$\varphi(i) \in \text{Com}(x^i) \quad \text{and} \quad x^{i+1} = x^{\varphi(i)}.$$

A path $C = (x^1, \ldots, x^r)$ is completely described by its starting point $x^1$ and the map $\varphi : \{1, \ldots, r-1\} \rightarrow \{1, \ldots, n\}$, called the strategy of the path [10,11], so by abuse of notation, we shall write $C = (x^1, \varphi)$, or $C = (x^1, (\varphi(1), \ldots, \varphi(r)))$ as well.

A path $(x^1, \ldots, x^r, x^1)$ with $r \geq 2$ is called a cycle (for $f$) and denoted $(x^1, \ldots, x^r)$. A cycle $C$ is completely described by one of its points, say $x^1$, and its strategy $\varphi : \{1, \ldots, r\} \rightarrow \{1, \ldots, n\}$, which enjoys $x^{r \varphi(r)} = x^1$, so we shall write $C = (x^1, \varphi)$ as well. Observe that a path $C = (x^1, \varphi)$ with strategy $\varphi : \{1, \ldots, r\} \rightarrow \{1, \ldots, n\}$ is a cycle if, and only if, for any $i = 1, \ldots, n$, the cardinality of $\varphi^{-1}(i)$ is even. As a consequence, $r$ is then even. A cycle $(x^1, (\varphi(1), \ldots, \varphi(r)))$ is said to be attractive when, for all $i = 1, \ldots, r$, $\text{Com}(x^i) = \{\varphi(i)\}$. Equivalently, for all $i = 1, \ldots, r$, $d_1(x^i, f(x^i)) = 1$, or

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

In other words, when you reach an attractive cycle, you cannot escape any more.
Given \( f : \{0, 1\}^n \to \{0, 1\}^n \), we attach to each \( x \in \{0, 1\}^n \) its \textit{discrete Jacobian matrix} \( J(x) \) as defined in [10,11]: \( J(x) \) is the \( n \times n \) matrix with \((i,j)\)-entry

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

Recall that a graph is said to be \textit{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 \( \{g_1, \ldots, g_n\} \) and an edge from \( g_j \) to \( g_i \) when \( J(x)_{i,j} = 1 \), i.e., the graph whose adjacency matrix if the transpose of \( J(x) \). M.-H. Shih and J.-L. Dong prove in [12] a discretised version of the Jacobian conjecture, which relates the dynamical behaviour of \( f \) with the graphs \( \Gamma(x) \): 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. On the other hand, Thomas' rules relate the dynamics of a population of genes to regulatory graphs, which are signed versions of the graphs \( \Gamma(x) \) and are defined now.

A \textit{regulatory graph} is a signed directed graph with vertex set \( \{g_1, \ldots, g_n\} \), i.e.,

\begin{align*}
\text{a directed graph with a sign, } +1 \text{ or } -1, \text{ attached to each edge.}
\end{align*}

\textbf{Definition 2.1} Given \( f : \{0, 1\}^n \to \{0, 1\}^n \) and \( x \in \{0, 1\}^n \), define \( G(x) \) to be the \textit{regulatory graph} 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.

Observe that \( \Gamma(x) \) is the graph underlying the regulatory graph \( G(x) \), and that our notion of regulatory graph is local, i.e., defined at each state \( x \), in a way analogous to [15] for instance. By enlarging the context, i.e., by taking unions of graphs on states \( x \), we can loose some information and recover more global notions of regulatory graphs as in [4,9,16]:

For any \( E \subseteq \{0, 1\}^n \), let

\[
G(E) = \bigcup_{x \in E} G(x)
\]

be the regulatory 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 the graph underlying \( G(E) \) is not necessarily simple (it can contain both a positive edge and a negative edge from one vertex to another).

A \textit{circuit} in a graph or in a regulatory graph \( G \) is a non-empty sequence \((g_{n_1}, \ldots, g_{n_k})\) of vertices such that \( G \) contains an edge from \( g_{n_i} \) to \( g_{n_{i+1}} \) for \( i = 1, \ldots, k-1 \) and an edge from \( g_{n_k} \) to \( g_{n_1} \). If \( I \subseteq \{1, \ldots, n\} \), an \( I \)-circuit is a circuit \((g_{n_1}, \ldots, g_{n_k})\) such that \( n_1, \ldots, n_k \in I \). If \( J \subseteq I \), a \( J \)-circuit is clearly an \( I \)-circuit. The \textit{sign} of a circuit \( C \) is the product of the signs of its edges.
A subcube \cite{12} is an affine subspace of \( \{0,1\}^n = (\mathbb{Z}/2\mathbb{Z})^n \), and for \( 0 \leq k \leq n \), a \( k \)-subcube is a \( k \)-dimensional subcube. For \( 0 \leq k \leq n \) and \( I \) a \( k \)-element subset of \( \{1, \ldots, n\} \), each \( x \in \{0,1\}^n \) generates a \( k \)-subcube \( x[I] = \{ y \in \{0,1\}^n \text{ s.t. } y_j = x_j \text{ for all } j \notin I \} \). Any subcube is clearly of the form \( x[I] \) for some \( x \) and some \( I \). A subcube of the form \( x[I] \) is called an \( I \)-subcube. If \( \kappa \) is an \( I \)-subcube, a \( \kappa \)-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 \( \kappa \), \( f \) has at most one \( \kappa \)-fixed point.

**Proof** Proceed by induction on the cardinality \( k \) of \( I \). The result holds for \( k = 0 \), since a 0-subcube is a singleton. If \( 0 \leq k \leq n - 1 \), 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 \( \kappa \)-fixed points \( a \) and \( b \) for some \( I \)-subcube \( \kappa \). There are two cases:

If \( d_1(a, b) = r \leq k \), there is an \( r \)-element set \( J \subseteq I \) such that \( b = \pi^J \), so \( a \) and \( b \) both belong the \( r \)-subcube \( \lambda = a[J] \) and are obviously \( \lambda \)-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 \( d_1(a, b) = k + 1 \), then \( b = \pi^I \). For each \( i \in I \), \( \overline{b}_i \in \kappa_i = a[I \setminus \{i\}] \), a \( k \)-subcube. Now, \( a \) is a \( \kappa_i \)-fixed point for each \( i \in I \). For each \( x \in \{0,1\}^n \), \( G(x) \) has no positive \((I \setminus \{i\})\)-circuit, so the induction hypothesis implies that \( \overline{b}_i \) is not a \( \kappa_i \)-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 \( \kappa \)-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 \cite{10,11}, this implies that \( G(b) \) has an \( I \)-circuit \( C = (g_{i_1}, \ldots, g_{i_t}) \) with \( i_1, \ldots, i_t \in I \) (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).

By hypothesis, \( C \) has to be negative, so we have:

\[
\begin{align*}
    b_{i_2} &= f_{i_2}(b) = \sigma_1(b_{i_1}) \\
    &\vdots \\
    b_{i_t} &= f_{i_t}(b) = \sigma_{t-1}(b_{i_{t-1}}) \\
    b_{i_1} &= f_{i_1}(b) = \sigma_t(b_{i_t}),
\end{align*}
\]

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

\[\square\]

As an immediate consequence, we have the following.

**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 $b = \pi^I$, 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. \[\square\]

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.

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 regulatory 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 regulatory graphs $G(x)$.
for \( x \) a state of \( C \) (Theorem 4.4), and that the attractiveness hypothesis is necessary. The stronger result that an attractive cycle implies the presence of a negative circuit in the regulatory graphs \( G(x) \) attached to a single state \( x \) (an optimal result) can be shown to hold for \( n \leq 3 \) by exhaustive exploration, but we have no proof of this conjecture for general \( n \).

### 4.1 Definitions and properties

If \( r \in \mathbb{N}, r \geq 1 \) 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[).\]

Let \( \sim \) be an equivalence relation on \( \{1, \ldots, r\} \).

**Definition 4.1** A bridge for \( \sim \) is a pair \((k, \ell)\) such that \( k \neq \ell \), \( k \sim \ell \) and for each \( i \in ]k, \ell[, k \neq i \).

If the equivalence class of \( k \in \{1, \ldots, r\} \) is not a singleton, then \( \sim \) has a unique bridge of the form \((k, \ell)\), which we denote by \((k, \sigma(k)) \); in particular, if all equivalence classes have cardinality at least 2, \( \sigma \) is a permutation of \( \{1, \ldots, r\} \). If \( X \) is a set and \( \varphi : \{1, \ldots, r\} \to X \), a bridge for \( \varphi \) is a bridge for the equivalence relation \( \sim_\varphi \) induced by \( \varphi \): \( k \sim_\varphi \ell \) if, and only if, \( \varphi(k) = \varphi(\ell) \).

**Lemma 4.2** If \( \sim \) is an equivalence relation on \( \{1, \ldots, r\} \) such that all equivalence classes have cardinality at least 2, then there exists a bridge \((k, \ell)\) for \( \varphi \) such that \( \{\varphi(1), \ldots, \varphi(\ell-1)\} = \{\varphi(1), \ldots, \varphi(r)\} \).

*Proof* Otherwise, for each \( k \in \{1, \ldots, r\} \), there is \( i \in \{\varphi(1), \ldots, \varphi(r)\} \) such that \([k, \sigma(k)] \cap \varphi^{-1}(i) = \emptyset \). By taking \( k' \in \varphi^{-1}(i) \) such that \([k', k] \cap \varphi^{-1}(i) = \emptyset \), we have \( \sigma(k') \notin [k, \sigma(k)[ \) because \( \sigma(k') \in \varphi^{-1}(i) \), therefore \([k, \sigma(k)[ \) is a strict subset of \([k', \sigma(k')[\). Hence, there is an infinite sequence \( k_0 = 1, k_1, k_2, \ldots \) such that \([k_i, \sigma(k_i)[ \subseteq [k_{i+1}, \sigma(k_{i+1})[ \) for each \( i \geq 0 \); see Figure 3. But then, \((k_i)_{i \geq 0} \) is an infinite sequence of distinct elements of \( \{1, \ldots, r\} \) and we have a contradiction. \( \square \)

### 4.2 Attractive cycles

**Lemma 4.3** Assume \( f : \{0, 1\}^n \to \{0, 1\}^n \) has an attractive cycle \((x^1, \ldots, x^r)\), with strategy \( \varphi : \{1, \ldots, r\} \to \{1, \ldots, n\} \). Then \( \varphi \) has at least one bridge, and for each bridge \((k, \ell)\) of \( \varphi \) and each \( i \in [k, \ell[, G(x^i) \) has an edge from \( g_{\varphi(i)} \) to \( g_{\varphi(i+1)} \) with sign \( \varepsilon_i \) such that \( \prod_{i \in [k, \ell[} \varepsilon_i = -1 \).

9
Fig. 3. Bridges in the proof of Lemma 4.2.

**Proof** Under the conditions of the Lemma, it is clear that the preimage of each element in the image of $\varphi$ has an even non-zero cardinality, so, for each $k = 1, \ldots, r$, $\varphi$ has a (unique) bridge $(k, \ell)$. Let then $(k, \ell)$ be a bridge of $\varphi$, and let $i \in [k, \ell]$. We take indices modulo $r$, i.e., we identify $r + j$ and $j$. Since the cycle $(x^1, \ldots, x^r)$ is asynchronous, we have:

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

so $x^{i+2}_{\varphi(i+1)} \neq x^{i+1}_{\varphi(i+1)}$. Since the cycle $(x^1, \ldots, x^r)$ is attractive, $f(x^i) = x^{i+1}$ and $f(x^{i+1}) = x^{i+2}$, hence:

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

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

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

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

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

Since in addition $(k, \ell)$ is a bridge, $x^\ell = x^k_{\varphi(k), \ldots, \varphi(\ell-1)}$ and we have $x^\ell_{\varphi(k)} = x^k_{\varphi(k)} \neq x^k_{\varphi(k)}$, so $\prod_{i \in [k, \ell]} \varepsilon_i = -1$. \[\square\]

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

10
Proof Since \((x^1, \ldots, x^r)\) is a cycle, the equivalence relation induced by \(\varphi\) enjoys the condition of Lemma 4.2, so \(\varphi\) has a bridge \((k, \ell)\) such that \(\{\varphi(k), \ldots, \varphi(\ell - 1)\} = \{\varphi(1), \ldots, \varphi(r)\}\). By Lemma 4.3, \((\varphi(k), \ldots, \varphi(\ell - 1))\) is a negative circuit of \(\bigcup_{i=1}^r G(x^i)\), and its vertex set is \(\{g_{\varphi(1)}, \ldots, g_{\varphi(r)}\}\). □

Let us remark that the choice of synchronous dynamics would have not led to a Thomas’ rule such as Theorem 4.4. Indeed, if we define a synchronous cycle for \(f : \{0, 1\}^n \to \{0, 1\}^n\) in the obvious way, i.e., to be a sequence \((x^1, \ldots, x^r)\) of states such that \(f(x^i) = x^{i+1}\) for \(i < r\) and \(f(x^r) = x^1\), the dynamics given in Figure 1 with \(n = 2\) has a synchronous cycle \(((0, 0), (1, 1))\), whereas the regulatory graph associated to any state has no negative circuit; of course there is no attractive asynchronous cycle.

4.3 Arbitrary cycles

The dynamics given in Figure 4 with \(n = 3\) (see [9]) has no attractive cycle, but has a non-attractive one:

\(\{(1, 0, 0), (1, 1, 0), (0, 1, 0), (0, 1, 1), (0, 0, 1), (1, 0, 1), (1, 0, 0)\}\)

and two fixed points \((0, 0, 0)\) and \((1, 1, 1)\). The regulatory graph associated to any state is a single positive circuit, so it has no negative circuit (despite the presence of the cycle), hence the attractiveness hypothesis is necessary to conclude with the presence of negative circuits. Yet, the following theorem asserts that the presence of a circuit, whatever its sign, follows from the existence of (non necessarily attractive) cycles.

**Theorem 4.5** If \(f : \{0, 1\}^n \to \{0, 1\}^n\) has a cycle \(C = (x^1, \ldots, x^r)\) with strategy \(\varphi\), then \(G(C) = G(x^1) \cup \cdots \cup G(x^r)\) has an \(I\)-circuit, where \(I\) is the image of \(\varphi\).

Proof Suppose that \(C = (x^1, \ldots, x^r) = (x^1, \varphi)\) is a cycle of minimal length, i.e., there is no cycle of length strictly smaller than \(r\). Let \(I \subseteq \{1, \ldots, n\}\) be the image of \(\varphi\). We want 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 \(x\), \(f_i(x) = f_i(\varphi')\).

Consider the equivalence class of \(j\), and \((k, l)\) a bridge for \(\sim_{\varphi}\). Denote \(x^k = x_{\varphi(1) \cdots \varphi(k-1)}\), \(x^k \in C\). The path \((x^k, \{\varphi(k+1), \ldots, \varphi(l-1)\})\) does exist (because \(f_i(x) = f_i(\varphi')\) for any \(i\) and \(x\)). Then, \(C' = (x^1, \{\varphi(1), \ldots, \varphi(k-1), \varphi(k+1), \ldots, \varphi(l-1), \varphi(l+1), \ldots, \varphi(r)\})\) is an cycle of length \(r - 2\): a contradiction. □
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 on the left are supposed to ease visualising the 3-cube.

5 Illustration and discussion

Figure 5 describes the Boolean asynchronous dynamics of a simple three-element genetic regulatory network. This dynamics has a single fixed point $(0,0,1)$ and an attractive cycle $((1,0,0), (1,0,1), (1,1,1), (1,1,0), (1,0,0))$, also denoted $((1,0,0), (3,2,3,2))$ (see Section 2.1).

On the basis of this dynamics, we can use the definitions introduced in this paper and provide some practical applications of our theorems. Let us compute the map $f$ and the discrete Jacobian matrix $J$ at the states $(0,0,0)$ and $(1,1,1)$.

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

The regulatory graph $G(0,0,0)$ thus contains two edges, one from $g_2$ to $g_3$, and a self-loop on gene $g_1$. As $x_1 = f_1(x)$ and $x_2 \neq f_2(x)$, the self-loop is
positive (activation) and gene $g_2$ is an inhibitor of gene $g_1$ (negative edge).

\[
\begin{array}{c}
G(0, 0, 0) \quad \text{is} \\
\begin{array}{c}
g_2 \\
g_3
\end{array}
\end{array}
\]

Let us turn now to state $(1, 1, 1)$:

\[
f(1, 1, 1) = (1, 1, 0) \quad \text{and} \quad J(1, 1, 1) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.
\]

The regulatory graph $G(1, 1, 1)$ contains four edges, a self-loop on $g_1$, as well as edges from $g_1$ to $g_2$, from $g_2$ to $g_3$ and from $g_3$ to $g_2$. As $x_1 = f_1(x)$, $x_1 = f_2(x)$, $x_2 \neq f_3(x)$ and $x_3 = f_2(x)$, all these interactions are positive (activations), except the edge from $g_2$ to $g_3$ which is negative.
In agreement with Theorem 4.4, the regulatory graph $G(1, 0, 0) \cup G(1, 0, 1) \cup G(1, 1, 1) \cup G(1, 1, 0)$ contains a negative circuit, and according to the strategy $(3, 2, 3, 2)$ of the attractive cycle in Figure 5, the vertices of this circuit are $g_2$ and $g_3$. The above computations show that $G(1, 1, 1)$ already has the negative circuit with vertices $g_2, g_3$.

In the regulatory graph associated to each of the two states just analysed, a positive self-loop on $g_1$ has been identified. The dynamics illustrated in Figure 5 does not exhibit several 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 to have a multistability, but rather an illustration of a tentative generalisation of the corresponding conjecture to the coexistence of alternative attractors, which still remain to be demonstrated.

The graph of Figure 5 encompasses one auto-activation on $g_1$, plus the negative circuit between $g_2$ and $g_3$, which is connected to the positive circuit through the positive edge from $g_1$ to $g_2$. The positive circuit is probably related to the separation of the state transition graph into two separated components, whereas the negative circuit is produced by the oscillatory behaviour of $g_2$ and $g_3$ for $x_1 = 1$.

From a biological point of view, it is interesting to note that similar negative circuits are found at the core of cell cycle and circadian rhythm controlling networks. Furthermore, in our example, the dynamical role of the negative circuit depends on the presence of product, thereby defining a functionality context, which corresponds to the biological notion of check point.

As the dimension $n$ increases, the inference of the regulatory graph $G(\{0, 1\}^n)$ requires the analysis of higher numbers of states to fully cover the corresponding Boolean space $\{0, 1\}^n$. However, for higher dimensions, it should be still possible to develop a systematic analysis of the Jacobian matrix around state subsets and induce the corresponding regulatory graphs, leading to the progressive delineation of the feedback circuits present in the original regulatory networks.

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) [7]. In this context, the main challenge consists in dealing with non-ideal time series, i.e., which do not include sufficiently high numbers of time points or sufficiently precise or reproducible measures. In this respect, note that our tentative approach focuses on structural properties of the network being inferred (presence of positive versus negative circuits, list of genes involved in these circuits), which could still largely hold even though
some intermediary elements acting in the original networks would be missed.

However, at this point, as our modeling approach does not take into account the time delays associated with each transition explicitly, the inferred regulatory graph will generally consist in simplifications of the underlying network, which could further include interactions involved in the adjustment of kinetic gene expression aspects.

References


