From Logical Regulatory Graphs to Standard Petri Nets: Dynamical Roles and Functionality of Feedback Circuits

Remy E., Ruet P. [1]

Institut de Mathématiques de Luminy - CNRS
Marseille, France

Mendoza L. [2]

Serono Pharmacological Research Institute
Geneva, Switzerland

Thieffry D., Chaouiya C. [3]

Laboratoire de Génétique et Physiologie du Développement
CNRS/Université de la Méditerranée - Marseille, France

Abstract

Logical modelling and Petri nets constitute two complementary approaches for the dynamical modelling of biological regulatory networks. On the one hand, the logical approach encompasses analytical tools enabling the delineation of specific relationships between the feedback structure and the dynamical properties of complex regulatory systems. On the other hand, the Petri net approach provides powerful algebraic methods and simulation tools. In this paper, we present a systematic way to translate logical models into standard Petri nets. We further propose a formalisation of the notion of circuit functionality in the Petri Net framework. Finally, we apply our approach to the modelling and the analysis of a molecular regulatory network involved in the control of Th-lymphocyte differentiation.

Key words: genetic regulatory graphs, Petri nets, feedback circuit, discrete dynamics, qualitative analysis.

[1] Email: {remy,ruet}@iml.univ-mrs.fr
[2] Email: Luis.Mendoza@serono.com
[3] Email: {chaouiya,thieffry}@lgpd.univ-mrs.fr

This is a preliminary version. The final version will be published in Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

1 Introduction

Regulatory networks are found at the core of all biological functions, from biochemical pathways, to gene regulation mechanisms, and cell communication processes. Their complexity often defies the intuition of the biologists and calls for the development of proper mathematical methods to model their structure and simulate their **dynamical behaviour**. A large variety of formal approaches have already been applied to biological regulatory networks, from ordinary or partial differential systems, to sets of stochastic equations (for a recent review, see [7]). However, as precise, quantitative information about the shape of regulatory functions or the values of involved parameters are generally lacking, **qualitative approaches** are usually more easily deployed.

One qualitative approach consists in modelling regulatory networks in terms of **logical equations**, using either Boolean or multi-level variables (see [8], [25] and references therein). The development of logical models for various biological networks already led to interesting insight into the network structures (in particular, regulatory feedback circuits) and the corresponding dynamical properties [25]. Relying on the generalised logical approach of R. Thomas, we have recently developed a software tool, GIN-sim, which enables the biologist to specify a regulatory model and check the qualitative evolution of the system for given initial states [6]. As the number of qualitative (*i.e.* logical or discrete) states grows exponentially with the number of elements involved in the regulatory network, there is a pressing need for proper analytical approaches to cope with the rapid delineation of larger regulatory networks.

At this stage, it appears interesting to articulate the logical method with another qualitative approach, namely the **Petri net** modelling [18,22]. Indeed, the Petri net formalism offers a framework to model, analyse and simulate the dynamical behaviour of large systems. As a first step in this direction, we have recently proposed a rigorous procedure to translate **logical regulatory models** into specific **regulatory Petri nets**, focusing on the Boolean case [5]. This formal bridge between these two approaches should help us to simultaneously exploit the corresponding analytical and simulation tools. Moreover, cross-fertilisation between these approaches should contribute to the delineation of new analytic procedures, *e.g.* the development of a feedback circuit analysis in the Petri net formalism, or the transposition of notions such as liveness in the logical framework. Finally, a proper combination of the two approaches should ease the modelling of mixed metabolic-genetic networks.

Most of the works related to Petri net approaches for the modelling of biological networks concentrate on quantitative aspects even though the kinetic parameters are generally not precisely accessible [9,13]. Based on a simulation approach, these works refer to several extensions of Petri nets, including hybrid PN, where places and transitions are either discrete, or continuous [2,16]. Qualitative aspects are discussed in [20,12]. One can identify clear correspondences between qualitative properties of the dynamics of biological networks and classical PN behavioural properties. For example, a dead marking represents a stable state of the system, while transition invariants correspond to cyclical trajectories.

Furthermore, Petri nets have been successfully applied to the modelling and the analysis of metabolic networks [20,13,15,11]. As emphasised in [28], one can draw extensive relationships between the traditional biochemical modelling and Petri net theory. In particular, the stoichiometry matrix of a metabolic network corresponds to the Petri net incidence matrix.

In the case of genetic regulatory networks, the correspondence is subtler because the semantics associated with the interactions between components varies. Indeed, regulators are usually not consumed during the regulatory processes. As far as we know, previous applications of Petri nets to genetic networks were written for particular systems (see *e.g.* [9,16]) and no systematic modelling procedure has been published for regulatory networks.

Amid the promising applications of Petri net theory to biological systems, we have proposed a systematic rewriting of Boolean regulatory graphs into a Petri net formalism in [5].

In the following section, we present our procedure to translate logical regulatory models into standard Petri net models. We further introduce a formal definition for the notion of regulatory feedback circuit **functionality** [25]. Next, these definitions are applied to a regulatory network involved in the control of Th-lymphocyte differentiation. Finally, conclusions and perspectives are proposed.

2 Boolean Regulatory Petri nets

In this section, we define the Petri net corresponding to a logical regulatory graph, *i.e.* whose dynamics simulates the dynamical behaviour of the underlying genetic regulatory network.

First, we briefly describe logical regulatory graphs, in the **Boolean** case, *i.e.* when the expression state of each gene takes its value in $\{0, 1\}$ (for more details on the formalism in the multivalued case, see [6]).

A **regulatory graph** is a labelled directed graph representing interactions between genes. Each interaction involves two genes (or other kinds of molecular components), the source and the target. As one gene can be the target of several interactions, we define, for each gene g_i, the subset $\mathcal{I}(i)$, called *input of* g_i, which contains the source genes of all incoming interactions on g_i. For each gene g_i, a **parameter** $K_i(X)$ is defined for each subset X of $\mathcal{I}(i)$. The value of this parameter determines the level to which g_i tends when X is the set of operating incoming interactions. As we consider here the Boolean case, these parameters take their values in $\{0,\ 1\}$. Summarising, a regulatory graph is defined by:

- a set of nodes $\mathcal{G} = \{g_1, \ldots, g_n\}$,

- a set $\mathcal{I}$ composed of n subsets $\mathcal{I}(i)$, defining the sources of interactions towards g_i, $\forall i \in \{1, \dots, n\}$,
- a set of parameters $\mathcal{K} = \{K_i(X), i = 1, \dots, n, X \subseteq \mathcal{I}(i)\}$.

2.1 Basic definitions and properties

Consider a Boolean regulatory graph $\mathcal{R} = (\mathcal{G}, \mathcal{I}, \mathcal{K})$. Let us define the following Petri net:

- To each gene correspond two (complementary) places g_i, $\overline{g_i}$, $i = \{1, \dots, n\}$, such that the sum of tokens in places g_i and $\overline{g_i}$ equals 1. The position of the token in g_i or $\overline{g_i}$ indicates whether the gene is ON or OFF. The set of places P thus contains $2\,n$ elements: $P = \mathcal{G} \cup \overline{\mathcal{G}}$, with $\overline{\mathcal{G}} \stackrel{\triangle}{=} \{\overline{g_1}, \dots, \overline{g_n}\}$.
- To each parameter $K_i(X)$, $i \in \{1, \dots, n\}, X \subseteq \mathcal{I}(i)$ corresponds a transition t^i_X. The transition t^i_X is enabled as soon as all places g_j of the set X AND all complementary places $\overline{g_j}$ of the set $(\mathcal{I}(i) \setminus X)$ (the complementary set of X in $\mathcal{I}(i)$) are marked.

Definition 2.1 Given a Boolean regulatory graph, $\mathcal{R} = (\mathcal{G}, \mathcal{I}, \mathcal{K})$, the associated **Boolean regulatory Petri net** (**BRPN**) $\mathbf{N}(\mathcal{R}) = (P, T, Pre, Post)$ is defined as follows:

- $P = \mathcal{G} \cup \overline{\mathcal{G}} = \{g_1, \overline{g_1}, \dots, g_n, \overline{g_n}\}$ is the set of places,
- $T = \{t^i_X,\ i = 1, \dots n,\ X \subseteq \mathcal{I}(i)\}$ is the set of transitions,
- $Pre : P \times T \to \{0, 1\}$ is the mapping defining edges between places and transitions (Pre-conditions).
- $Post : T \times P \to \{0, 1\}$ is the mapping defining edges between transitions and places (Post-conditions).

The functions Pre and $Post$ are defined as follows:

(i) Case $g_i \notin \mathcal{I}(i)$ (g_i is not a self-regulator). For a given transition t^i_X, the only terms to be defined (*i.e.* all the other terms are equal to zero) are:

$$Pre(g_i, t^i_X) = Post(t^i_X, \overline{g_i}) = 1 - K_i(X)\,, \tag{1}$$
$$Pre(\overline{g_i}, t^i_X) = Post(t^i_X, g_i) = K_i(X)\,, \tag{2}$$
$$Pre(g_j, t^i_X) = Post(t^i_X, g_j) = 1 \qquad \forall g_j \in X\,, \tag{3}$$
$$Pre(\overline{g_j}, t^i_X) = Post(t^i_X, \overline{g_j}) = 1 \qquad \forall g_j \in \mathcal{I}(i) \setminus X\,. \tag{4}$$

(ii) Case $g_i \in \mathcal{I}(i)$ (g_i is a self-regulator). Consider a given transition t^i_X,

- if $g_i \in X$, the only case to be considered is $K_i(X) = 0$ (cf. Remark 2.2). Therefore, the only terms to be defined are:

$$Pre(g_i, t^i_X) = Post(t^i_X, \overline{g_i}) = 1\,, \tag{5}$$
$$Pre(g_j, t^i_X) = Post(t^i_X, g_j) = 1 \qquad \forall g_j \in X\,, g_j \neq g_i\,, \tag{6}$$
$$Pre(\overline{g_j}, t^i_X) = Post(t^i_X, \overline{g_j}) = 1 \qquad \forall g_j \in \mathcal{I}(i) \setminus X\,. \tag{7}$$

- if $g_i \notin X$, the only case to be considered is $K_i(X) = 1$. Therefore, the only terms to be defined are:

$$Pre(\overline{g_i}, t^i_X) = Post(t^i_X, g_i) = 1\,, \tag{8}$$
$$Pre(g_j, t^i_X) = Post(t^i_X, g_j) = 1 \qquad \forall g_j \in X\,, \tag{9}$$
$$Pre(\overline{g_j}, t^i_X) = Post(t^i_X, \overline{g_j}) = 1 \qquad \forall g_j \in \mathcal{I}(i) \setminus X\,, g_j \neq g_i\,. \tag{10}$$

Comments: Equations (1)-(2) state that if the parameter $K_i(X)$ equals 1, g_i is an input and $\overline{g_i}$ an output of the corresponding transition t^i_X. In other words, there can be a decrease of the level of g_i if it is already present. Symmetrically, if $K_i(X) = 0$, then $\overline{g_i}$ is an input and g_i is an output of the corresponding transition t^i_X. Equations (3)-(4), (6)-(7) and (9)-(10) state that the elements contributing to the combination of interactions involved in $K_i(X)$ (*i.e.* which are in X) are the inputs of the corresponding transitions and are not consumed by these transitions.

Remark 2.2 Note that the case of a self-regulation is slightly different. Let g_i be a self-regulator; there should be one transition t^i_X for each parameter $K_i(X)$. But two cases do not lead to any change on g_i: when $g_i \notin X$ (*i.e.* g_i is absent) and $K_i(X) = 0$, and when $g_i \in X$ (*i.e.* g_i is present) and $K_i(X) = 1$. The resulting transitions are never enabled.

Thus, two representations of self-regulations are possible: an explicit one, where all the four cases are considered, or a compact one, where the two cases described above are omitted, because they do not have any influence on the dynamics of the net. We choose here to use the compact representation. Consequently,

- when $g_i \in X$ and $K_i(X) = 0$, g_i is only an input of transition t^i_X, $\overline{g_i}$ is an output of t^i_X,
- when $g_i \notin X$ and $K_i(X) = 1$, $\overline{g_i}$ is only an input of transition t^i_X, and g_j an output of t^i_X.

Figure 1 illustrates the BRPNs corresponding to an activation (left) and to an inhibition (right) between two genes, respectively (to simplify the notation, we have denoted $K_i(\{g_j\})$ by $K_i(g_j)$ and $t^i_{\{g_j,\ldots,g_k\}}$ by $t^i_{j,\ldots,k}$).

Although the graphical representation of a BRPN is more complex than the corresponding regulatory graph (indeed the BRPN represents the regulatory graph together with its parameterisation), the analytical tools provided by the Petri nets should be usefull in our context.

Notation 1 *Let denote* $\widehat{K}_i(X) \triangleq 2\,K_i(X) - 1$, *and* d_i *the number of parameters for* g_i *(for all* $i \in \{1, \ldots n\}$*). If* $g_i \notin \mathcal{I}(i)$*, then* $d_i = 2^{\#\mathcal{I}(i)}$*; otherwise,* g_i *is self-regulated, and* d_i *is smaller and depends on the combined effect of incoming interactions together with the self-regulation (cf. Case 2 in Definition 2.1). We will denote* X^j_i *the* j*th subset of* $\mathcal{I}(i)$*. Then, the* **incidence matrix** $C \triangleq Post^T - Pre$ *is a* $2n \times (\sum_{i=1\ldots n} d_i)$ *matrix. Its components take their values in* $\{-1, 0, 1\}$*, and* C *has the following structure:*

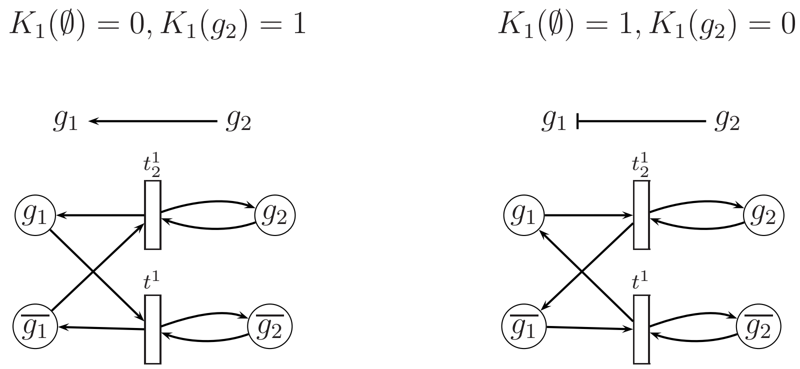

Fig. 1. Regulatory Petri nets corresponding to the activation of g_1 by g_2 (left) and to the inhibition of g_1 by g_2 (right). Top: regulatory graph representation and logical parameters (K's) for each type of regulation. $K_1(\emptyset)$ represents the basal expression of the first gene, *i.e.* in the absence of the regulatory product of the gene 2, whereas $K_1(g_2)$ represents the expression of that gene in the presence of the regulatory product. Bottom: corresponding Petri net representation, with two places for each gene. Transition t^1 corresponds to parameter $K_1(\emptyset)$, while t_2^1 corresponds to $K_1(g_2)$.

$$C = \begin{pmatrix} \boxed{D_1} & \mathit{0} & \dots & \mathit{0} \\ \mathit{0} & \boxed{D_2} & \dots & \mathit{0} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \mathit{0} & \mathit{0} & \dots & \boxed{D_n} \end{pmatrix} \quad \mathit{with} \quad D_i \triangleq \begin{pmatrix} \widehat{K}_i(X_i^1) & \dots & \widehat{K}_i(X_i^{d_i}) \\ -\widehat{K}_i(X_i^1) & \dots & -\widehat{K}_i(X_i^{d_i}) \end{pmatrix}.$$

Definition 2.3 Given a regulatory Petri net $\mathbf{N}(\mathcal{R}) = (P, T, Pre, Post)$, a **valid marking** $M : P \rightarrow \{0, 1\}$ corresponds to a state of the Boolean regulatory graph $\mathcal{R}$ and verifies: $\forall g_i \in \mathcal{G}$, $M(g_i) = 1 - M(\overline{g_i})$.

In the sequel, we will only consider valid markings.

The following property shows the equivalence between the asynchronous state transition graph related to the Boolean regulatory graph and the reachability graph of the corresponding BRPN. A state S in the state transition graph is described by a valid marking, *i.e.*, the current values associated to each element (1 if the regulatory product g_i is present at this state, 0 otherwise). Note that we only consider the fully asynchronous assumption, as it is generally the case in the Petri net framework. We could also consider the fully synchronous assumption, modifying the firing rule (in this case, all enabled transitions would fire simultaneously).

Property 1 *There exists an edge from state S_1 to state S_2 in the asynchronous dynamical graph related to a Boolean regulatory graph $\mathcal{R} = (\mathcal{G}, \mathcal{K})$, iff there exists an enabled transition t in the associated BRPN such that M_1*

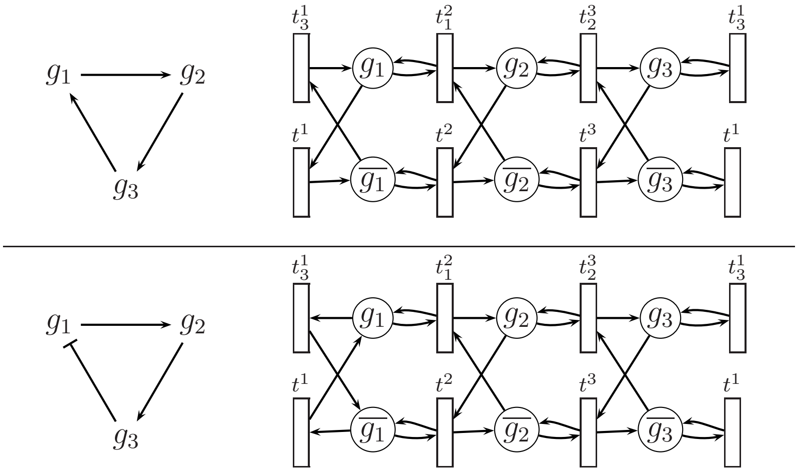

Fig. 2. Three-element regulatory circuits and the corresponding Petri nets. Upper row: positive circuit. Lower row: negative circuit. Note that transitions t^1 and t_3^1 are repeated, illustrating the general structure of the net.

verifies $M_1[t\rangle M_2$ *(*t *is enabled by* M_1 *and its firing leads to the marking* M_2*) with, for all* $i = 1, \ldots n$*,*

$$M_1(g_i) = S_1(i) \quad M_1(\overline{g_i}) = 1 - S_1(i)\,,$$
$$M_2(g_i) = S_2(i) \quad M_2(\overline{g_i}) = 1 - S_2(i)\,.$$

2.2 Isolated regulatory circuits

In the context of the logical approach, several authors have emphasised the crucial dynamical roles of regulatory circuits. More specifically, **positive** circuits (*i.e.* containing an even number of inhibitions) have been associated to multistationarity, whereas **negative** circuits (with an odd number of inhibitions) can generate sustained periodic behaviour [25]. In what follows, we derive a general Petri net formulation for these two classes of circuits and check their dynamical properties.

In the case of isolated regulatory circuits, each gene g_i is the target of a unique interaction exerted by g_{i-1}, and is the source of a unique interaction towards g_{i+1} (here and in the sequel, indices are considered *modulo* n, the length of the circuit, *i.e.* $i + n = i$).

Let $\mathcal{C} = (\mathcal{G}, \mathcal{K})$ be a regulatory circuit, with $\mathcal{G} = \{g_1, \ldots, g_n\}$ and $\mathcal{K} = \{K_i(\emptyset),\, K_i(g_{i-1})\}_{i=1,\ldots,n}$, recalling that $\mathcal{I}(i) = \{g_{i-1}\}$ (cf. [21] for more details). In this simpler context, when we consider the interaction from g_i to g_{i+1}, the values of parameters for which the circuit is *functional* (see Section 2.3) are $K_{i+1}(\emptyset) = 0$ and $K_{i+1}(g_i) = 1$ (we say that this interaction is an activation), or $K_{i+1}(\emptyset) = 1$ and $K_{i+1}(g_i) = 0$ (the interaction is an inhibition).

As we shall see, the corresponding regulatory Petri nets $\mathbf{N}(\mathcal{C})$ also have a well defined structure. Property 2 directly follows from Definition 2.1 and Notation 1 (cf. [5] for the proof):

Property 2 *Let* $\mathbf{N}(\mathcal{C}) = (\mathcal{G} \cup \overline{\mathcal{G}}, T, Pre, Post)$ *be the regulatory Petri net corresponding to the Boolean regulatory circuit* $\mathcal{C}$*. The incidence matrix* C *is a* $2n \times 2n$ *block-diagonal matrix, where diagonal blocks* D_i *are the* 2×2 *submatrices defined by :*

$$D_i = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \qquad or \qquad D_i = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

depending on whether the interaction of g_{i-1} *on* g_i *is an activation or an inhibition, respectively.*

In [21], we have proved that an isolated positive circuit generates two stable states, which are *mirroring* each other (a component is "on" in one state iff it is "off" in the other state), and that an isolated functional negative circuit leads to a dynamical graph where all states feed a specific dynamical circuit of length twice the number of elements in the circuit. These results can be restated in the Petri net formalism (see also [5]):

Property 3 • *Let* $\mathbf{N}(\mathcal{R})$ *be a regulatory Petri net corresponding to an isolated (functional) positive regulatory circuit, then there are exactly two dead valid markings* M_d^1 *and* M_d^2 *which are mirroring each other. Each of these two markings is reachable from any live valid marking.*

- *Let* $\mathbf{N}(\mathcal{R})$ *be a regulatory Petri net corresponding to a (functional) negative regulatory circuit and* E *be the set of all valid markings which enable exactly one transition.*

(i) *Any valid marking of the net is live.*
(ii) E *has* $2n$ *elements and is organised as a cycle.*
(iii) *Each marking in* E *is reachable from any valid marking.*

2.3 Functionality

In the previous section, we have described the case of isolated circuits, showing that their dynamical behaviours depend on their signs. More generally, for complex regulatory networks, R. Thomas has enounced rules binding the dynamical behaviour to the presence of specific circuits. More precisely, he conjectured that a necessary condition for multistationarity is the presence of a **positive circuit**, whereas a necessary condition for homeostasis and/or sustained, stable oscillations is the presence of a **negative circuit**. These rules have already been formally stated and partly demonstrated in the differential formalism [23,10,24], but not yet in the discrete modelling, except within a restricted Boolean additive case [3].

The main point in formulating Thomas' rules is to take into account the notion of *functionality*: the rough intuition is that a circuit would be functional when the conjectured behaviour is observable at some state(s) of the biological system. To illustrate this idea, consider the work in [24]: the dynamics is governed there by an ordinary differential equation $\dot{x} = f(x)$, where $x \in \mathbb{R}^n$ represents the concentrations and f is a regular enough map from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then, a regulatory graph G_x is attached to each state x by means of the Jacobian matrix of f at x: a family of regulatory graphs is defined on the state space, and this dependency between a graph and a state (locality) can be viewed as a way of restricting attention to those interactions which are functional at that state.

In this section, we propose a formal definition of the notion of the functionality of an interaction and, more generally, of a circuit. This definition is local, *i.e.*, functionality is defined *within a specific context*, since the functionality of a circuit embedded in a more complex regulatory network may depend on the effect of some combination of incoming interactions. An illustration of these notions can be found in Section 3. The definitions are expressed in the Petri net framework, but they can be easily adapted to other discrete models.

Definition 2.4 Functionality of an interaction Let $\mathcal{R} = (\mathcal{G}, \mathcal{K})$ be a Boolean regulatory graph, and $\mathbf{N}(\mathcal{R}) = (P, T, Pre, Post)$ its associated Boolean regulatory Petri net. Let us consider an interaction $I : g_i \longrightarrow g_j$ from the graph $\mathcal{R}$, and X, X' a partition of $\mathcal{I}(j) \setminus \{g_i\}$. When

$$Post(t^j_X, g_j) \neq Post(t^j_{X\cup\{g_i\}}, g_j)\,,$$

we say that I is **functional in the context** (X, X'). Then, its **functionality marking set** is denoted $\mathcal{S}_I \subseteq \{M : P \to \{0,1\}\}$, with

$$\mathcal{S}_I = \{M : P \to \{0,1\} \ \text{ s.t. } \ M(g) = 1 \ \forall g \in X\,, \ \text{ and } \ M(g) = 0 \ \forall g \in X'\}\,.$$

Therefore, an interaction is said to be functional within some context, if the modification of the level of expression of its source leads to a change of the level of the target. The functionality of a circuit will then depends on the existence of a context for which all the transitions of this circuit are functional.

Definition 2.5 Functionality of a circuit A marking M belongs to the functionality marking set of a circuit $\mathcal{C}$ if it belongs to the intersection $\mathcal{S}_\mathcal{C}$ of the functionality marking sets of all interactions of $\mathcal{C}$.

If such a marking exists ($\mathcal{S}_\mathcal{C} \neq \emptyset$), then $\mathcal{S}_\mathcal{C}$ is the functionality marking set of the circuit $\mathcal{C}$.

Note that if $\mathcal{S}_\mathcal{C} = \emptyset$, we say that the circuit is not functional.

We propose an algorithm to determine the functionality marking set of a given circuit $\mathcal{C}$ embedded in a regulatory graph $\mathcal{R}$. Let $\mathbf{N}(\mathcal{R}) = (P, T, Pre, Post)$ be the BRPN associated to $\mathcal{R}$. Assume that $\mathcal{C}$ consists in n interactions $I^i : g_i \to g_{i+1}$, for $i = 1, \ldots, n$ (notation *modulo* n). For each i, we apply the following procedure:

(i) Determine the set of transitions $\mathrm{Trans}(I^i)$ associated to I^i:
$\mathrm{Trans}(I^i) = \{t \in T, C(g_{i+1}, t) = 1 \text{ or } C(\overline{g_{i+1}}, t) = 1\}$, where C is the incidence matrix of $\mathbf{N}(\mathcal{R})$ (it is the set of t_X^{i+1}, for all $X \subseteq \mathcal{I}(i+1)$).

(ii) Determine the pairs of transitions $(t, t') \in \mathrm{Trans}(I^i)$ having the same inputs except g_i and g_{i+1}: $Pre(g_i, t) \neq Pre(g_i, t')$, and for all g different from g_i, g_{i+1}, $Pre(g, t) = Pre(g, t')$.

(iii) For each such pair (t, t'), test whether the interaction is functional:
$Post(t, g_{i+1}) \neq Post(t', g_{i+1})$. Then, for each such (t, t'),
- The context of functionality (X, X') is given by:
$X = \{g \in \mathcal{I}(i+1) \setminus \{g_i\} \text{ s.t. } Pre(g, t) = Pre(g, t') = 1\}$
$X' = \{g \in \mathcal{I}(i+1) \setminus \{g_i\} \text{ s.t. } Pre(g, t) = Pre(g, t') = 0\}$.
- We define the set $\mathcal{S}_{t,t'}$ of markings M such that for each $g \in X$, $M(g) = 1$ and for each $g \in X'$, $M(g) = 0$.

(iv) Define $\mathcal{S}_i$ to be the union of all the $\mathcal{S}_{t,t'}$.

Finally, the functionality marking set of $\mathcal{C}$ is the intersection of the functionality marking sets $\mathcal{S}_i$, $i = 1, \ldots, n$.

3 Application to the modelling of Th-lymphocyte differentiation

3.1 Introducing Th-cell differentiation

The vertebrate immune system contains diverse cell populations, like antigen presenting cells, natural killer cells, and B and T lymphocytes. Among the latter, CD4+ T helper lymphocytes, upon receiving an appropriate antigenic stimulus, can further differentiate into T-helper 1 (Th1) or Th2 cells, which enable cell mediated immunity and humoral responses, respectively. Th1 and Th2 cells can be distinguished according to their pattern of cytokine secretion. Immune responses biased towards the Th1 phenotype result in autoimmune diseases, while enhanced Th2 responses originate allergic reactions [1,19]. Various mathematical models have been proposed for the differentiation, activation and proliferation of Th-lymphocytes. Up to now, many of these models were focusing on interactions between immunological cell populations at a macroscopic level [4,27]. Other model analyses aim at understanding the mechanism of the generation of antibody and T-cell receptors diversity, as well as the dynamical properties of the large networks defined by the interactions between cytokines [14] or between immunoglobulins (see *e.g.* [26]).

In order to illustrate our Petri net approach to the analysis of genetic regulatory networks, we present the translation of a simplified Boolean model describing the interactions between the most prominent lymphokines, lymphokine receptors and transcription factors which control Th-cell differentiation following activation by antigen recognition (see Figure 3). The biological justification of the set of interactions encompassed by this model can be found in [17], accompanied with an extensive list of references to the immunological literature.

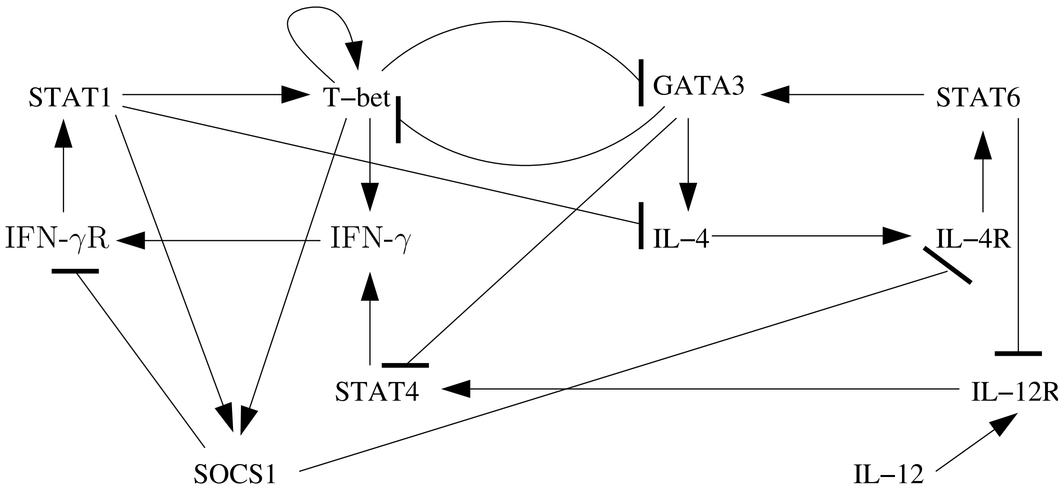

Fig. 3. Regulatory graph of the network controlling Th lymphocyte differentiation. The nodes represent transcription regulatory factors (T-bet, GATA-3), signaling transduction factors (STAT1, STAT4, STAT6, SOCS-1), lymphokines (IFN-γ, IL-4, IL-12) and receptors (IFN-γR, IL-4R, IL-12R), whereas the edges represent activations (arrows) or inhibitions (T ends) between these components.

Here, we focus on the delineation of the multi-stability properties of the network as a result of the functionality of specific positive regulatory circuits found in the corresponding regulatory graph. Furthermore, we show that this type of feedback circuit analysis can be processed directly in the context of the Petri net formalism, relying on the sole analysis of the Pre and $Post$ matrices.

Given the regulatory graph in Figure 3, and the set of parameters in Table 1, we construct the associated BRPN $(P, T, Pre, Post)$ with:

- the set of places $P = \{g_1, \overline{g_1}, \dots g_{12}, \overline{g_{12}}\}$ (the numbering of the places is defined in Table 1),
- the set of transitions $T = \{t^i_X, i = 1 \dots 12, X \subset \mathcal{I}(i)\}$ (where t^i_X denotes the transition corresponding to the parameter $K_i(X)$).

The complete definition of the BRPN corresponding to the regulatory graph is defined through the matrices Pre (of size (24×39)) and $Post$ (of size (39×24)) (not shown).

The regulatory graph of Figure 3 contains 18 regulatory circuits (15 positive, three negative), involving from one to ten elements. Using the logical formalism, it can be shown that nine of these circuits (seven positive, two negative) are functional. Still in the context of the logical approach, three stable states are found [17]. In what follows, we show that these three stable states actually correspond to dead markings in the corresponding BRPN, and we apply the procedure described in Section 2.3 to one particular circuit.

Genes		Non-zero parameters	Stable states		
			S_1	S_2	S_3
IFN-γ	(1)	$K_1(9)$ $K_1(11)$ $K_1(9,11)$	0	0	1
IL-4	(2)	$K_2(12)$	0	1	0
IL-12	(3)		0	0	0
IFN-γR	(4)	$K_4(1)$	0	0	0
IL-4R	(5)	$K_5(2)$	0	1	0
IL-12R	(6)	$K_6(3)$	0	0	0
STAT1	(7)	$K_7(4)$	0	0	0
STAT6	(8)	$K_8(5)$	0	1	0
STAT4	(9)	$K_9(6)$	0	0	0
SOCS1	(10)	$K_{10}(7)$ $K_{10}(11)$ $K_{10}(7,11)$	0	0	1
T-bet	(11)	$K_{11}(7)$ $K_{11}(11)$ $K_{11}(7,11)$	0	0	1
GATA-3	(12)	$K_{12}(8)$	0	1	0

Table 1

For each gene of the model presented in Figure 3: the non-zero parameters (44 parameters are defined in all; no parameter are associated with $IL-12$ as it acts as an input of the system), and the corresponding expression level in each stable state (S_1, S_2, S_3) are specified.

3.2 Stable states and their biological interpretation

It has been shown [17] that the system can reach the three stable states of Table 1. The first stable state (S_1) corresponds to the virgin Th cells, whereas the second and third stable states correspond to Th2 and Th1 differentiated lymphocytes, respectively. We can easily check that these 3 stable states correspond to dead markings for the corresponding BRPN, verifying that: $\forall t \in T, \exists p \in P \, \text{s.t.} \, Pre(p,t) = 1 \, \text{and} \, M(p) = 0$. We will only give here some flavour of the proof:

- For S_1, where all genes are OFF ($S_1(g_i) = 0$, $i = 1, \ldots 12$), it is clear that no transition is enabled because of the matrix Pre which verifies: $\forall t \in T, \exists g \, \text{s.t.} \, Pre(g,t) = 1$.
- For S_2, $S_2(g_2) = S_2(g_4) = S_2(g_8) = S_2(g_{12}) = 1$ whereas $S_2(g_i) = 0$ for the other genes. $S_2(g_2) = 1$ could enable t^2, t^2_7, $t^2_{7,12}$, t^5_2 and $t^5_{2,10}$, but:
  $Pre(\overline{g_{12}}, t^2) = Pre(\overline{g_{12}}, t^2_7) = 1$ and $S_2(g_{12}) = 1$ ($\Leftrightarrow S_2(\overline{g_{12}}) = 0$),
  $Pre(g_7, t^2_7) = 1$ and $S_2(g_7) = 0$,
  $Pre(g_5, t^5_2) = 1$ and $S_2(g_5) = 0$,
  $Pre(g_{10}, t^5_{2,10}) = 1$ and $S_2(g_{10}) = 0$.

Consequently, no transition is enabled under the marking S_2.

A similar reasoning can be developed for S_3. However, a more extensive dynamical analysis would require the investigation of the stable state reachability, *i.e.* the verification of the existence of a path leading from specific initial conditions to these stable states.

3.3 Regulatory circuits: dynamical roles and functionality constraints

To illustrate our approach, let us analyse the following positive circuit: $\mathcal{C}$= [IL-4, IL-4R, STAT6, GATA-3]. In order to determine the functionality marking set of this circuit, we begin by applying the first step of the procedure proposed in Section 2.3, on each of the four interactions defining $\mathcal{C}$, successively. This leads to the selection of the following submatrices of *Pre* (Table 3) and *Post* (Table 2) .

	t^2	t^2_7	t^2_{12}	$t^2_{7,12}$	t^5	t^5_2	t^5_{10}	$t^5_{2,10}$	t^8	t^8_5	t^{12}	t^{12}_8	t^{12}_{11}	$t^{12}_{8,11}$
IL-4 (2)	0	0	1	0	0	1	0	1	0	0	0	0	0	0
	1	1	0	1	1	0	1	0	0	0	0	0	0	0
IL-4R (5)	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	0	0	0	0	1	0	1	1	1	0	0	0	0	0
STAT1 (7)	0	1	0	1	0	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0	0	0
STAT6 (8)	0	0	0	0	0	0	0	0	0	1	0	1	0	1
	0	0	0	0	0	0	0	0	1	0	1	0	1	0
SOCS1 (10)	0	0	0	0	0	0	1	1	0	0	0	0	0	0
	0	0	0	0	1	1	0	0	0	0	0	0	0	0
T-bet (11)	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	0	0	0	0	0	0	0	0	0	0	1	1	0	0
GATA-3 (12)	0	0	1	1	0	0	0	0	0	0	0	1	0	0
	1	1	0	0	0	0	0	0	0	0	1	0	1	1

Table 2
Sub-matrix $Post^T$ for the circuit $\mathcal{C}$ =[IL-4, IL-4R, STAT6, GATA-3] and its inputs.

Note that the set of genes to be considered encompasses not only the four genes of $\mathcal{C}$ (IL-4, IL-4R, STAT6, GATA-3), but also the genes exerting inputs on these genes (STAT1, SOCS1, T-bet).

Following the procedure, we now need to apply steps 2-4 to each interaction. We limit ourselves here to their applications to the interaction from IL-4 (2) to IL-4R (5). The second step of the procedure leads to the selection of the pairs (t^5, t^5_2) and $(t^5_{10}, t^5_{2,10})$. Then, proceeding with step 3:

- For the pair (t^5, t^5_2), we have $Post(t^5, g_5) \neq Post(t^5_2, g_5)$, and the interaction is thus functional in the context $X = \emptyset$ and $X' = \{g_{10}\}$. The corresponding functionality marking set is $\mathcal{S}_2 = \{M : P \to \{0, 1\} \text{ s.t. } M(g_{10}) = 0\}$.
- For the second pair, we have $Post(t^5_{10}, g_5) = Post(t^5_{2,10}, g_5)$. Consequently,

	t^2	t^2_7	t^2_{12}	$t^2_{7,12}$	t^5	t^5_2	t^5_{10}	$t^5_{2,10}$	t^8	t^8_5	t^{12}	t^{12}_8	t^{12}_{11}	$t^{12}_{8,11}$
IL-4 (2)	1	1	0	1	0	1	0	1	0	0	0	0	0	0
	0	0	1	0	1	0	1	0	0	0	0	0	0	0
IL-4R (5)	0	0	0	0	1	0	1	1	0	1	0	0	0	0
	0	0	0	0	0	1	0	0	1	0	0	0	0	0
STAT1 (7)	0	1	0	1	0	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0	0	0
STAT6 (8)	0	0	0	0	0	0	0	0	1	0	0	1	0	1
	0	0	0	0	0	0	0	0	0	1	1	0	1	0
SOCS1 (10)	0	0	0	0	0	0	1	1	0	0	0	0	0	0
	0	0	0	0	1	1	0	0	0	0	0	0	0	0
T-bet (11)	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	0	0	0	0	0	0	0	0	0	0	1	1	0	0
GATA-3 (12)	0	0	1	1	0	0	0	0	0	0	1	0	1	1
	1	1	0	0	0	0	0	0	0	0	0	1	0	0

Table 3
Sub-matrix *Pre* for the circuit $\mathcal{C}$ =[IL-4, IL-4R, STAT6, GATA-3] and its inputs.

these transitions add no further marking to the current functionality marking set.

The same reasoning can be followed for the three other interactions of the circuit $\mathcal{C}$, leading to: $\mathcal{S}_5 = \{M : P \to \{0,1\}\}$ (all possible valid markings), $\mathcal{S}_8 = \{M : P \to \{0,1\}$ s.t. $M(g_{11}) = 0\}$, and $\mathcal{S}_{12} = \{M : P \to \{0,1\}$ s.t. $M(g_7) = 0\}$.

Finally, the functionality marking set of the circuit $\mathcal{C}$ is the intersection of the four resulting marking sets: $\mathcal{S}_\mathcal{C} = \mathcal{S}_2 \cap \mathcal{S}_5 \cap \mathcal{S}_8 \cap \mathcal{S}_{12} = \{M : P \to \{0,1\}$ s.t. $M(g_7) = M(g_{10}) = M(g_{11}) = 0\}$.

In conclusion, the positive circuit $\mathcal{C}$ is functional provided that the products STAT1 (g_7), SOCS1 (g_{10}) and T-bet (g_{11}) are absent. It is interesting to note that, if we switch the value of the parameter $K_2\{12\}$ to zero, this positive circuit looses its functionality. This loss of functionality results in the loss of the stable state Th2 (S_2).

4 Conclusions and prospects

In this paper, we have described a systematic way to translate logical models of genetic regulatory networks into standard Petri nets (see also [5]). We have particularly focused on the delineation of a procedure to determine the marking sets insuring the functionality of the feedback circuits found in genetic regulatory graphs. This procedure leads to interesting insights into the temporal behaviour of the modelled system, as it is well established that regulatory circuits are found at the origin of fundamental dynamical properties, such as multistability or homeostasis. The functionality marking set of a circuit defines the sub-region of the phase space where this circuit generates the corresponding dynamical property (*e.g.* the presence of a separatrix in the positive circuit case).

To illustrate our approach, we have delineated the Petri net translation of a Boolean model for the molecular regulatory network controlling Th-lymphocyte differentiation [17]. On the basis of this translation, we have shown that all three logical stable states correspond to dead markings in the Petri net model. These states correspond to specific lymphocyte populations: the virgin Th population, and the differentiated Th1 *versus* Th2 populations, corresponding respectively to the enhancement of cellular *versus* humoral immune response. Focusing on one functional positive circuit, we have further illustrated how our formal procedure can be applied to our Th-lymphocyte network model in order to delineate the corresponding functionality marking set. Moreover, we have shown that the contradiction of the functionality constraints (*e.g.* by changing the value of a logical parameter in the logical formalism, or the topology of the incidence matrix in the Petri net model) leads to the loss of one of the stable states.

In order to generalise this approach, we now need to develop software tools to automate the evaluation of the number and the identity of all dead markings in Petri net regulatory models. In addition, we need to address the question of their reachability from specific initial markings. There exist already several softwares which perform this kind of analysis. But the characteristic structures of the regulatory Petri nets should allow the development of improved algorithms.

To cover the full expression power of the generalised logical formalism [25], we are currently working on the generalisation of our rewriting rules to encompass multilevel logical models.

Logical modelling and Petri nets constitute two complementary approaches for the dynamical modelling of biological regulatory networks. The combination of these approaches opens new prospects for the analysis of complex regulatory networks. In particular, using our rewriting rules for genetic regulatory interactions should ease the modelling and analysis of mixed metabolic-genetic networks.

References

[1] Agnello D., C.S.R. Lankford, J. Bream, A. Morinobu, M. Gadina, J. O'Shea and D.M. Frucht, *Cytokines and transcription factors that regulate T helper cell differentiation: new players and new insights*, J. Clin. Immunol. **23** (2003), 147-161.

[2] Alla H. and R. David,*Continuous and hybrid petri nets*, Journal of Circuits, Systems, and Computers **8-1** (1998), 159–188.

[3] Aracena J., "Modèles mathématiques discrets associés à des systèmes biologiques. Application aux réseaux de régulation génétiques," Thèse de

Doctorat Spécialité Mathématiques Appliquées, Université Joseph Fourier, Grenoble, 2001.

[4] Bergmann C., J.L. Van Hemmen and L.A. Segel, *Th1 or Th2: how an appropriate T helper response can be made*, Bull. Math. Biol. **63** (2001), 405–430.

[5] Chaouiya C., E. Remy, P. Ruet and D. Thieffry, *Qualitative Modelling of Genetic Networks: From Logical Regulatory Graphs to Standard Petri Nets*, J. Cortadella and W. Reisig (Eds), ICATPN04, LNCS **3099** (2004), 137–156.

[6] Chaouiya C., E. Remy, B. Mossé and D. Thieffry, *Qualitative analysis of regulatory graphs: a computational tool based on a discrete formal framework*, L. Benvenuti, A. De Santis, L. Farina (Eds), POSTA03, LNCIS **294** (2003), 119–126.

[7] de Jong H., *Modeling and Simulation of Genetic Regulatory Systems: A Literature Review*, J. Comp. Biol. **9** (2002), 67–103.

[8] Glass L. and S.A. Kauffman, emphThe logical analysis of continuous, non-linear biochemical control networks, J.Theor.Biol. **39** (1973), 103–129.

[9] Goss P.J.E. and J. Peccoud, *Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets*, Proc. Natl. Acad. Sci. USA. **95** (1998), 6750–6755.

[10] Gouzé J.L., *Positive and negative circuits in dynamical systems*, J. Biol. Syst. **6** (1998): 11–15

[11] Heiner M., I. Koch and K. Voss, *Analysis and Simulation of Steady States in Metabolic Pathways with Petri Nets*, K. Jensen (Ed.), CPN'01-Third Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools. University of Aarhus, Denmark, (2001),15–34. (2001).

[12] Heiner M. and I. Koch, *Petri Net Based Model Validation in Systems Biology*, J. Cortadella and W. Reisig (Eds), ICATPN04, LNCS **3099** (2004), 216–237.

[13] Hofestädt R. and S. Thelen, *Quantitative Modeling of Biochemical Networks*, In Silico Biology **1** (1998), 39-53.

[14] Krueger G.R.H., G.R. Marshall, U. Junker, H. Schroeder, L.M. Buja and G. Wang, *Growth factors, cytokines, chemokines and neuropeptides in the modeling of T-cells*, In vivo **16** (2002), 365-386.

[15] Küffner R., R. Zimmer and T. Lengauer, *Pathway analysis in metabolic databases via differential metabolic display (DMD)*, Bioinformatics **9** (2000), 925-936.

[16] Matsuno H., Y. Tanaka, H. Aoshima, A. Doi, M. Matsui and S. Miyano, *Biopathways representation and simulation on hybrid functional Petri net*, In Silico Biology **3** (2003), 0032,
URL: http://www.bioinfo.de/isb/toc_vol_03.html.

[17] Mendoza L., *A network model for the control of the differentiation process in Th cells*, Submitted.

[18] Murata T., *Petri Nets: Properties, Analysis and Applications*, Proceedings of the IEEE **77** (1989), 541–580.

[19] Murphy K.M. and S.L. Reiner, *The lineage decisions on helper T cells*, Nat. Rev. Immunol. **2** (2002), 933-944.

[20] Reddy V.N., M.N. Liebman and M.L. Mavrovouniotis, *Qualitative analysis of biochemical reaction systems*, Comput. Biol. Med. **26** (1996), 9-24.

[21] Remy E., B. Mossé, C. Chaouiya and D. Thieffry, *A description of dynamical graphs associated to elementary regulatory circuits*, Bioinformatics **19** (2003), ii172–178.

[22] Reisig W., "Petri Nets", Springer-Verlag (1985).

[23] Snoussi E.H., *Necessary conditions for multistationatity and stable periodicity*, J. Biol. Syst. **6** (1998),3–9

[24] Soulé C., *Graphic requirements for multistationarity*, ComPlexUs **1** (2003),123-133.

[25] Thomas R., D. Thieffry and 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.

[26] Weisbuch G., R.J. De Boer, A.S. Perelson, *Localized memories in idiotypic networks*, J. theor. Biol. **146** (1990), 483–499.

[27] Yates A., C. Bergman, J.L. Van Hemmen, J. Stark and, R. Callard, *Cytokine-mediated regulation of helper T cell populations*, J. theor. Biol. **206** (2000), 539–560.

[28] Zevedei-Oancea I. and S. Schuster, *Topological analysis of metabolic networks based on Petri net theory*, In Silico Biology **3** (2003), 0029,
URL: http://www.bioinfo.de/isb/toc_vol_03.html.