



# Influence Networks compared with Reaction Networks: Semantics, Expressivity and Attractors

François Fages, Thierry Martinez, David A Rosenblueth, Sylvain Soliman

## ► To cite this version:

François Fages, Thierry Martinez, David A Rosenblueth, Sylvain Soliman. Influence Networks compared with Reaction Networks: Semantics, Expressivity and Attractors. 2017. hal-01510216v1

**HAL Id: hal-01510216**

**<https://inria.hal.science/hal-01510216v1>**

Preprint submitted on 19 Apr 2017 (v1), last revised 14 Feb 2018 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Influence vs Reaction Systems: Semantics, Expressivity and Attractors

François Fages, Thierry Martinez, David A. Rosenblueth and Sylvain Soliman

**Abstract**—In Systems Biology, modelers develop more and more reaction-based models to describe the mechanistic biochemical reactions underlying cell processes. They may also work, however, with a simpler formalism of influence graphs, to merely describe the positive and negative influences between molecular species. The first approach is promoted by reaction model exchange formats such as SBML, while the second is supported by QualSBML and tools that have been historically developed to reason about boolean gene regulatory networks. In practice, modelers often reason with both kinds of formalisms, and may find an influence model useful in the process of building a reaction model. In this paper, we introduce a formalism of influence systems with forces, and put it in parallel with reaction systems with rates, in order to develop a similar hierarchy of boolean, Petri Net, stochastic and differential semantics. We show that the expressive power of influence systems is the same as that of reaction systems under the differential semantics, but weaker under the other interpretations, in the sense that some discrete behaviours of reaction systems cannot be expressed by influence systems. This approach leads us to consider a positive boolean semantics, which we compare to the asynchronous semantics with negation of gene regulatory networks *à la* Thomas. We study the monotonicity properties of the positive semantics and derive from them an algorithm to compute attractors in both the positive and negative semantics. This algorithm is evaluated on models of the literature about the p53/Mdm2 DNA damage repair system, the circadian clock, and on a challenging influence model of the MAPK signaling network.

**Index Terms**—Reaction systems, Influence systems, Differential Equations, Stochastic semantics, Petri Nets, Boolean semantics, Galois connections, Attractors.



## 1 INTRODUCTION

IN SYSTEMS BIOLOGY modelers develop more and more reaction models to describe the biochemical reactions underlying cell processes. This approach is promoted by reaction-model exchange formats such as SBML [1] and by the subsequent creation of large reaction-based model repositories such as BioModels [2], without prejudging of their interpretation by differential equations, continuous-time Markov chains, Petri nets, or boolean transition systems [3].

Modelers can also work, however, with a simpler formalism of *influence systems* to merely describe the positive and negative influences between molecular species, without fixing their implementation with biochemical reactions. In particular, boolean influence systems have been popularized in the 70's by Glass, Kauffman [4] and Thomas [5], [6] to reason about gene regulatory networks, represented by ordinary graphs between genes given with a boolean transition table which defines their boolean transition semantics, be it synchronous or asynchronous. QualSBML is an SBML package dedicated to this kind of system. Necessary conditions for multi-stability (cell differentiation) and oscillations (homeostasis) have been given in terms of positive or negative circuits in the influence graph [7], [8]. Several tools such as GINsim [9], GNA [10] or Griffin [11], use these properties and powerful graph-theoretic and model checking tech-

niques to automate reasoning about the boolean state transition graph, compute attractors and verify various reachability and path properties. The representation of boolean influence systems by Petri nets was described in [12] but leads to complicated encodings. It is also worth mentioning that influence systems with spatial information have been nicely developed in [13] as a formalism particularly suitable for describing natural algorithms in life sciences and social dynamics.

In Systems Biology, modelers often reason with both kinds of formalisms, and may find an influence model useful to consider and maintain in the process of building a reaction model, for instance in order to reduce it while preserving the essential influence circuits [14]. One reason is that it is easier to visualize influence systems, rather than reaction systems for which sophisticated graphical conventions such as SBGN [15] have been developed. While it is clear that the influence graph is an abstraction of the reaction hypergraph [3], and perhaps more surprisingly that the influence system derived from the signs of the Jacobian matrix of the differential semantics of a reaction system is essentially independent of the rate functions [16], influence models are mostly used for their graphical representation and their boolean semantics, but more rarely as a modeling paradigm for systems biology with quantitative semantics using differential equations, or stochastic processes.

In this paper, we introduce a formalism of influence systems with forces, which we put in parallel with reaction systems with rates, in order to develop a similar hierarchy of boolean, Petri Net, stochastic and differential semantics for influence systems [3], in the framework of abstract interpretation originally introduced for reasoning

- François Fages and Sylvain Soliman are with Inria, University Paris-Saclay, France  
Contact: see <http://lifeware.inria.fr/>
- Thierry Martinez is with Inria Paris, France
- David A. Rosenblueth is with Universidad Nacional Autónoma de México, Mexico

Manuscript received April 15, 2017.

about programs [17]. We show that the expressive power of influence systems is the same as that of reaction systems under the differential semantics, but is weaker under the other interpretations, in the sense that some formal discrete behaviours of reaction systems cannot be expressed by influence systems. This approach provides an influence model with a hierarchy of possible interpretations related by precise abstraction relationships, so that, for instance, if a behaviour is impossible in the boolean semantics, it is surely not possible in the stochastic semantics whatever the influence forces are.

This leads us to consider a positive boolean semantics which we compare to the asynchronous semantics with negation of gene regulatory networks *à la* Thomas. We study the monotonicity properties of the positive boolean semantics and derive from them an algorithm to compute attractors in both the positive and negative semantics. These concepts are illustrated, and the attractor computation algorithm evaluated, on models of the p53/Mdm2 DNA damage repair system [18], [19], [20], of the circadian clock [21], and a challenging 53 variable model of the influence of MAPK signaling on cell-fate decision in urinary bladder cancer [22].

For the sake of reproducibility, all the examples presented in this paper can be run in Biocham v4 notebooks online<sup>1</sup>.

## 2 PRELIMINARIES ON REACTION SYSTEMS

### 2.1 Notations

In this article, unless explicitly noted, we will denote sets and multisets by capital letters (e.g.  $S$ , we shall also use calligraphic letters for some sets), tuples of values by vectors (e.g.,  $\vec{x}$ ), and elements of those sets or vectors (e.g. real numbers, functions) by small Roman or Greek letters. For a multiset  $M : S \rightarrow \mathbb{N}$ ,  $M(x)$  denotes the multiplicity of element  $x$  in  $M$  (usually the stoichiometry). By abuse of notation,  $\geq$  will denote the integer or boolean pointwise order for vectors, multisets and sets (i.e. set inclusion), and  $+$ ,  $-$  the corresponding operations for adding or removing elements. With these unifying notations, set inclusion may thus be noted  $S \leq S'$  and set difference  $S - S'$ .

### 2.2 Syntax

We recall here definitions from [16], [23] for directed reactions with inhibitors<sup>2</sup>. In this paper, we assume a finite set  $S = \{x_1, \dots, x_s\}$  of molecular species.

**Definition 1.** A reaction over  $S$  is a quadruple  $(R, I, P, f)$ , also noted  $f$  for  $R/I \Rightarrow P$ , where

- $R$  is a multiset of reactants in  $S$ ,
- $I$  a multiset of inhibitors in  $S$ ,
- $P$  a multiset of products in  $S$ ,

- and  $f : \mathbb{R}^s \rightarrow \mathbb{R}$  is a rate function over molecular species (concentrations or numbers).

A reaction system is a finite set of reactions.

It is worth noting that a molecular species in a reaction can be both a reactant and a product, i.e. a *catalyst*, or both a reactant and an inhibitor (e.g. Botts–Morales enzymes [24]). Catalysts and inhibitors are not distinguished in SBML where both are called reaction *modifiers*. Moreover unlike SBML, we find it useful to consider only directed reactions (reversible reactions being represented here by two reactions) and to enforce the following compatibility conditions between the rate function and the structure of a reaction:

**Definition 2** ([16], [23]). A reaction  $(R, I, P, f)$  over molecular species  $\{x_1, \dots, x_s\}$  is well-formed if the following conditions hold:

- 1)  $f(x_1, \dots, x_s)$  is a partially differentiable function, non-negative on  $\mathbb{R}_+^s$ ,
- 2)  $x_i \in R$  if and only if  $\partial f / \partial x_i(\vec{x}) > 0$  for some value  $\vec{x} \in \mathbb{R}_+^s$ ,
- 3)  $x_i \in I$  if and only if  $\partial f / \partial x_i(\vec{x}) < 0$  for some value  $\vec{x} \in \mathbb{R}_+^s$ ,
- 4)  $x_i \in R$  implies  $f(x_1, \dots, x_s) = 0$  whenever  $x_i = 0$ .

A reaction system is well-formed if all its reactions are well-formed. Those conditions ensure that the reactants contribute positively to the rate of the reaction at least in some region of the concentration space (condition 2), that the inhibitors indeed decrease the reaction rate (condition 3), and that the system remains positive (condition 4) [23].

**Example 1.** The classical birth-death process model of Lotka–Volterra can be represented by the following well-formed reaction system (without reaction inhibitors) between a proliferating prey  $A$  and a predator  $B$ :

$k1 * A * B \text{ for } A+B \Rightarrow 2*B.$   
 $k2 * A \text{ for } A \Rightarrow 2*A.$   
 $k3 * B \text{ for } B \Rightarrow \_.$

For an example of reaction with inhibitor, one could consider the specific inhibition of the proliferation rate of  $A$  by some other variable  $C$  (e.g. competing for nutrients) as follows:

$k2 * A / (1+C) \text{ for } A / C \Rightarrow 2*A.$

It is worth noticing that such an inhibited reaction shows a negative influence of  $C$  on  $A$ , as would also have a degradation of  $A$  by  $C$  for instance, but with a fundamentally different mechanism.

**Example 2.** For a “real” example of reaction with inhibitor from the literature, one can cite the model of the DNA damage repair system by Ciliberto et al. [18], on which we will return in Sec. 5.1. This ODE model contains a term which represents the phosphorylation, inhibited by the p53 protein, of the Mdm2 protein. This can be represented by the following well-formed reaction with inhibitor:

$k * \text{Mdm2} / (j + p53) \text{ for } \text{Mdm2} / p53 \Rightarrow \text{Mdm2p}.$

1. <http://lifeware.inria.fr/biocham4>, see the notebooks in <http://lifeware.inria.fr/wiki/software/#TCBB17>

2. One technical difference with respect to [16], [23] is that we consider here reactions with a multiset of inhibitors instead of a set. The reason comes from the discrete semantics with negation which was not considered before, and in which the multiplicity of the inhibitors can then be interpreted as threshold values above which a reaction does not proceed.

### 2.3 Hierarchy of Semantics

As detailed in [3], a reaction system can be interpreted with different formalisms that are formally related by abstraction relationships in the framework of abstract interpretation [17] and form a hierarchy of semantics. We simply recall here the definitions of the different semantics of a reaction system.

The *differential semantics* associates an Ordinary Differential Equation (ODE) system to a reaction system  $R$  in the usual way:

$$\frac{dx_j}{dt} = \sum_{(R_i, I_i, P_i, f_i) \in \mathcal{R}} (P_i(j) - R_i(j)) \times f_i$$

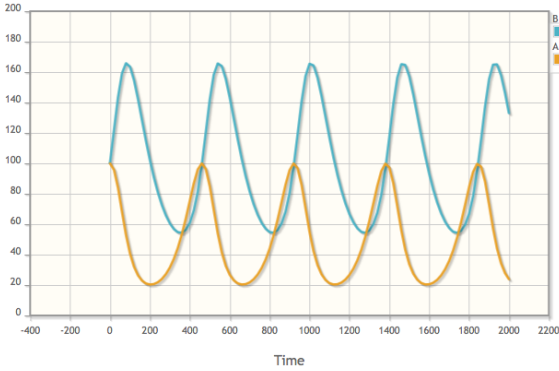
It is worth noting that in this interpretation, the inhibitors are supposed to decrease the reaction rate but do not prevent the reaction from proceeding with effects on the products and reactants.

**Example 3.** In Ex. 1, the differential semantics gives the classical Lotka–Volterra equations

$$dA/dt = k_2 * A - k_1 * A * B$$

$$dB/dt = k_1 * A * B - k_3 * B$$

and the well-known oscillations between the concentrations of the prey and the predator:

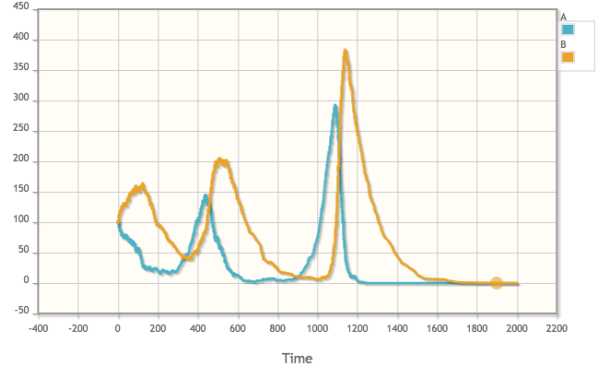


The *stochastic semantics* for reaction systems is defined from a transition relation, noted  $\rightarrow_S$ , between discrete states describing numbers of each molecule, i.e. vectors  $\vec{x}$  of  $\mathbb{N}^s$ . A transition is enabled if there are enough reactants to fire a reaction. Each reaction is associated with a transition relation with propensity given by the rate function as follows:

$$\forall (R_i, I_i, P_i, f_i), \vec{x} \xrightarrow{f_i}_S \vec{x}' \text{ with propensity } f_i \\ \text{if } \vec{x} \geq R_i \text{ and } \vec{x}' = \vec{x} - R_i + P_i$$

Transition probabilities between discrete states are obtained through normalization of the propensities of all enabled reactions, and the time of next reaction can be computed from the rates *à la* Gillespie [25]. In this interpretation, the inhibitors are supposed to decrease the reaction propensity but do not prevent the reaction from occurring. They are thus ignored here by the stochastic transition conditions as in the differential semantics.

**Example 4.** In Ex. 1, the stochastic interpretation can exhibit some oscillations similar to the differential interpretation, but almost surely the extinction of the predator, i.e. a qualitatively different behaviour:



The *Petri Net semantics* is defined similarly on discrete states with a transition relation  $\rightarrow_D$  which just ignores the rate functions and is thus a trivial abstraction of the stochastic semantics by a forgetful functor:

$$\forall (R_i, I_i, P_i, f_i), \vec{x} \rightarrow_D \vec{x}' \text{ if } \vec{x} \geq R_i, \vec{x}' = \vec{x} - R_i + P_i$$

The *boolean semantics* is similar to the *Petri Net* one but on boolean vectors  $x$  of  $\mathbb{B}^s$ , obtained by the “zero, non-zero” abstraction of integers. With this abstraction, when the number of a molecule is decremented, it can still remain present, or become absent. It is thus necessary to take into account all the possible complete consumption or not of the reactants in order to obtain a correct boolean abstraction of the Petri Net and stochastic semantics [3]. The *boolean transition system*  $\rightarrow_B$  is thus defined by:

$$\forall (R_i, I_i, P_i, f_i), \forall S \leq R_i, \vec{x} \rightarrow_B \vec{x}' \\ \text{if } \vec{x} \geq R_i, \vec{x}' = \vec{x} - S + P_i.$$

The boolean semantics provides an over-approximation of the qualitative behaviours of the stochastic semantics. In particular, the reachability of steady states (at which the system can stay forever but with a possibility to escape) and stable states (that the system cannot leave) can be verified by CTL model checking [26]. As proven in [3], the discrete semantics are related by successive Galois connections, which means that if a behaviour is not possible in the boolean semantics, it is not possible in the stochastic semantics whatever the reaction rates are.

**Example 5.** In Ex. 1, the boolean semantics reveals, by CTL model checking, the possible qualitative behaviours of the discrete semantics including the stochastic semantics, i.e. either the steady but not stable persistence of the predator (which can always disappear if the prey becomes extinct), the stable persistence of the prey (if the predator becomes extinct), or the extinction of both the predator and the prey, and the absence of boolean oscillations:

```
biocham: present ({A,B}).
biocham: generate_ctl_not.
reachable (steady (B))
reachable (stable (A))
reachable (stable (not A))
reachable (stable (not B))
```

On the other hand, the differential semantics is generally not an abstraction of the stochastic behaviour, but when all the numbers of molecules tend to the infinity, an approximation of the mean stochastic behaviour, as shown for instance in [25]. In Ex. 1, these limit conditions are obviously not

satisfied and the mean stochastic behaviour is very different from the differential behaviour. This approximation result nevertheless motivates the interpretation of inhibitors as acting solely on the reaction rates, and not as negative conditions for the firing of reactions in the discrete semantics.

## 2.4 Discrete Semantics with Negation

Though not common for reaction systems, contrarily to influence systems studied in the next section, one can also consider discrete semantics with negation where the set of inhibitors of a reaction is seen as a conjunction of negative conditions for the transition (disjunctions can be represented with several reactions). In order to express negative conditions on multisets, let us denote by  $<_0$  the strict pointwise order between two multisets restricted to the intersection of their supports.

The *boolean with negation transition system*  $\longrightarrow_{BN}$  can then be defined by:

$$\forall (R_i, I_i, P_i, f_i) \forall S \leq R_i, \vec{x} \longrightarrow_{BN} \vec{x}' \text{ if } \vec{x} \geq R_i, \vec{x} <_0 I_i, \vec{x}' = \vec{x} - S + P_i$$

and the *stochastic and Petri net semantics with negation* can be defined similarly just by adding the condition  $\vec{x} <_0 I_i$  to the transitions.

It is worth remarking that in the Petri Net semantics with negation, the strict interpretation of the reaction inhibitors by negation makes it possible to express *bounded multilevel semantics* which enjoys many decidability properties, by interpreting the stoichiometric coefficients as activation levels, and by adding an inhibitor of the form  $k * x$  with level coefficient bound  $k$  to the reactions that produce  $x$ .

## 2.5 Influence Graph of a Reaction System

Here we recall two definitions of the influence graph<sup>3</sup> associated with a reaction system, and their equivalence under general assumptions [16], [23]. The first definition is based on the Jacobian matrix  $J$  formed of the partial derivatives  $J_{ij} = \partial \dot{x}_i / \partial x_j$ , where  $\dot{x}_i$  is defined by the differential semantics.

**Definition 3.** The differential influence graph (DIG) associated with a reaction system is the graph having for vertices the molecular species, and for edge-set the following two kinds of edges:

$$\{A \rightarrow^+ B \mid \partial \dot{x}_B / \partial x_A > 0 \text{ for some value } \vec{x} \in \mathbb{R}_+^s\} \\ \cup \{A \rightarrow^- B \mid \partial \dot{x}_B / \partial x_A < 0 \text{ for some value } \vec{x} \in \mathbb{R}_+^s\}$$

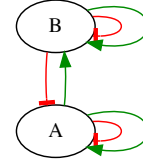
**Definition 4.** The syntactical influence graph (SIG) associated with a reaction system  $M$  is the graph having for vertices the molecular species, and for edges the following set of positive and negative influences:

$$\{A \rightarrow^+ B \mid \exists (R_i, I_i, P_i, f_i) \in M \\ (R_i(A) > 0 \text{ and } P_i(B) - R_i(B) > 0) \\ \text{or } (I_i(A) > 0 \text{ and } P_i(B) - R_i(B) < 0)\} \\ \cup \{A \rightarrow^- B \mid \exists (R_i, I_i, P_i, f_i) \in M \\ (R_i(A) > 0 \text{ and } P_i(B) - R_i(B) < 0) \\ \text{or } (I_i(A) > 0 \text{ and } P_i(B) - R_i(B) > 0)\}$$

The syntactical graph is trivial to compute, in linear time, by browsing the syntax of the rules.

3. By a slight abuse of terminology, we call influence graph in fact a labeled multigraph, since there may be both a positive and a negative influence from one vertex to another one.

**Example 6.** In the Lotka–Volterra reaction system, both definitions of the influence graph give the same graph:



The equality always holds if the syntactical influence graph contains no conflict, i.e. no pair of the form  $A \rightarrow^+ B$  and  $A \rightarrow^- B$  between the same molecules:

**Theorem 1** ([16], [23]). For any well-formed reaction system (resp. such that the syntactical influence graph contains no conflict), the differential influence graphs is included in (resp. identical to) the syntactical influence graph.

## 3 INFLUENCE SYSTEMS WITH FORCES

In this section, we propose a syntax for influence systems with forces which makes it possible to develop influence models of biochemical processes within a hierarchy of differential, stochastic, Petri Net and boolean semantics, similarly to reaction systems. We then compare their expressive power to reaction systems, and focus on different variants of their discrete semantics to compare them to Thomas's setting for gene regulatory networks.

### 3.1 Syntax

The idea to disambiguate the semantics of the different influences incoming on a vertex of the influence graph, is to syntactically distinguish conjunctive from disjunctive conditions in an influence system, by writing influences with several sources for representing a conjunction of conditions, while the different influences on a same target express a disjunction of conditions. These syntactical conventions are a particular case of the concept of multiplexes introduced in [27] restricted here to disjunctive normal forms.

Given  $S = \{x_1, \dots, x_s\}$  a set of species, a

**Definition 5.** An influence system  $I$  is a set of quintuples  $(P, I, t, \sigma, f)$  called influences, where

- $P$  is a multiset on  $S$ , called positive sources of the influence,
- $I$  a multiset of negative sources,
- $t \in S$  is the target,
- $\sigma \in \{+, -\}$  is the sign of the influence, accordingly called either positive or negative influence,
- and  $f$  is a real-valued mathematical function of  $\mathbb{R}^s$ , called the force of the influence.

Influences of sign  $+$  are called positive influences and those of sign  $-$ , negative influences. In addition, we distinguish the positive sources from the negative sources in an influence (positive or negative), in order to annotate the fact that in the differential semantics, the source increases or decreases the force of the influence, and in the boolean

semantics with negation whether the source or the negation of the source is a condition for a change in the target.

In the examples below, we use the ASCII syntax of Biocham v4 for influences. Positive (resp. negative) influences are written with an arrow  $\rightarrow$  (resp.  $\leftarrow$ ) which separates the sources from the target. The positive and negative sources are separated by a  $/$ , which can be omitted if there is no negative source.

**Example 7.** *The birth-death model of Lotka–Volterra of Ex. 1 can also be represented by the following well-formed influence system:*

```
k1*A*B for A,B -< A.
k1*A*B for A,B -> B.
k2*A for A->A.
k3*B for B-<B.
```

composed of four well-formed influences, with positive sources only,  $(\{A, B\}, \emptyset, A, -, k1*A*B)$ ,  $(\{A, B\}, \emptyset, B, +, k1*A*B)$ ,  $(\{A\}, \emptyset, A, +, k2*A)$  and  $(\{B\}, \emptyset, B, -, k3*B)$ . These four influences, two of which containing a conjunction of two sources, are a rewriting of the six edges of the influence graph depicted in Ex. 6.

For an example of influence with inhibitor, one can consider the specific inhibition of the proliferation rate of  $A$  by some variable  $C$  which is distinguished from a general negative influence of  $C$  on  $A$ , by writing  $C$  as an inhibitor of the positive influence of  $A$  on  $A$ :

```
k2*A/(1+C) for A/C -> A.
```

Let us now define the concept of *well-formed* influence systems similarly to the above Def. 2 for reaction systems.

**Definition 6.** *An influence  $(P, I, t, \sigma, f)$  over molecular species  $\{x_1, \dots, x_s\}$  is well-formed if the following conditions hold:*

- 1)  $f(x_1, \dots, x_s)$  is a partially differentiable function, non-negative on  $\mathbb{R}_+^s$ ;
- 2)  $x_i \in P$  if and only if  $\sigma = +$  (resp.  $-$ ) and  $\partial f / \partial x_i(\vec{x}) > 0$  (resp.  $< 0$ ) for some value  $\vec{x} \in \mathbb{R}_+^s$ ;
- 3)  $x_i \in I$  if and only if  $\sigma = +$  (resp.  $-$ ) and  $\partial f / \partial x_i(\vec{x}) < 0$  (resp.  $> 0$ ) for some value  $\vec{x} \in \mathbb{R}_+^s$ ;
- 4)  $t \in P$  if  $\sigma = -$ .

### 3.2 Hierarchy of Semantics

The *differential semantics* of an influence system  $\mathcal{I}$  over  $S$ , is given by (the solutions of) the following ODE system:

$$\frac{dx_k}{dt} = \sum_{(P_i, I_i, x_k, +, f_i) \in \mathcal{I}} f_i - \sum_{(P_j, I_j, x_k, -, f_j) \in \mathcal{I}} f_j$$

This ODE adds up all the forces of the positive influences on  $x_k$  and subtracts all forces of negative influences on  $x_k$  in the derivative of  $x_k$  over time. In Ex. 7, one can easily check that this definition leads to the same equations as for Ex. 1.

The negative sources in a well-formed influence decrease the force of the influence but do not disable it. Consequently, we define similarly the *stochastic semantics* of an influence system with forces, by a transition system, noted  $\rightarrow_S$ , between discrete states, i.e. vectors  $\vec{x}$  of  $\mathbb{I}^s$ , with the condition that the positive sources are present in sufficient number and no condition on the negative sources, the transition

propensity is defined by the force, and the target is updated as follows:

$$\forall (P_i, I_i, t_i, \sigma_i, f_i), \vec{x} \xrightarrow{f_i}_S \vec{x}' \text{ with propensity } f_i \\ \text{if } \vec{x} \geq P_i \text{ and } \vec{x}' = \vec{x} \sigma_i t_i$$

Transition probabilities between discrete states are obtained through normalization of the propensities of all enabled transitions, and the time of next reaction can also be given *à la* Gillespie [25]. In this interpretation, the negative sources are supposed to decrease the influence propensity but do not prevent the influence from proceeding. They are thus ignored here by the stochastic transition conditions.

The *Petri Net* (PN) semantics simply ignores the forces:  $\forall (P_i, I_i, t_i, \sigma_i, f_i), \vec{x} \xrightarrow{D} \vec{x}'$  if  $\vec{x} \geq P_i, \vec{x}' = \vec{x} \sigma_i t_i$ . Note that since a target has no multiplicity coefficient, a PN transition always adds or subtracts one to the level of the target, and cannot define a self-loop in the state transition graph.

The *positive boolean semantics* is defined on boolean vectors  $x$  of  $\mathbb{B}^s$ , by the “zero, non-zero” abstraction of integers of the Petri Net semantics. Consequently, the boolean semantics associates two transitions to one negative influence: one that inactivates the target, and one that loops with the target active<sup>4</sup>:

$$\forall (P_i, I_i, t_i, +, f_i), \vec{x} \xrightarrow{B} \vec{x}' \text{ if } \vec{x} \geq P_i, \vec{x}' = \vec{x} + t_i \\ \forall (P_i, I_i, t_i, -, f_i), \vec{x} \xrightarrow{B} \vec{x}' \text{ if } \vec{x} \geq P_i, \vec{x}' = \vec{x} \text{ or } \vec{x}' = \vec{x} - t_i$$

This boolean semantics is positive in the sense that it ignores the negative sources of an influence and contains no negation in the enabling condition. In Ex. 7, one can check that the boolean transitions are the same as in Ex. 1, although in general one can expect to get more transitions (Thm. 3.3 and Prop. 1 below).

With these definitions, we obtain as in [3] a hierarchy of semantics related by simple abstraction relationships in the framework of abstract interpretation [17], i.e. Galois connections between a concrete domain and an abstract domain, noted  $(\mathcal{C}, \sqsubseteq_{\mathcal{C}}) \xleftrightarrow{\alpha}_{\gamma} (\mathcal{A}, \sqsubseteq_{\mathcal{A}})$ , which will be here limited to the discrete semantics domains formalized by set lattices ordered by inclusion :

**Theorem 2.** *For an influence system  $\mathcal{I}$ , the stochastic semantics, the Petri Net semantics and the boolean semantics are related by Galois connections,  $(\{\vec{x} \xrightarrow{f_i}_S \vec{x}'\}, \subseteq) \xleftrightarrow{\alpha_{SD}}_{\gamma_{SD}} (\{\vec{x} \xrightarrow{D} \vec{x}'\}, \subseteq) \xleftrightarrow{\alpha_{DB}}_{\gamma_{DB}} (\{\vec{x} \xrightarrow{B} \vec{x}'\}, \subseteq)$*

*Proof.* Let us first remark that our transition semantics are sets and that in powerset domains, the pointwise extension of any function from the base set of the concrete domain to the abstract domain forms a Galois connection [3]. Thus we just have to define the abstraction functions pointwise to obtain Galois connections with respect to the set inclusion ordering.

The first abstraction of the stochastic semantics by the Petri Net semantics is a trivial abstraction by a forgetful functor  $\alpha_{SD}$ , namely the erasing of the propensities in the transitions.

The second abstraction  $\alpha_{DB}$  abstracts discrete states by boolean states. The domain of boolean transitions is related to the Petri Net transitions semantics domain by the

4. The loop condition was not mentioned in [28].

zero/non-zero abstraction from the integers to the booleans,  $\alpha_{NB} : \mathbb{N} \rightarrow \mathbb{B}$ , and its pointwise extension from discrete states to boolean states.

We can easily check that the boolean transitions are indeed the abstraction by  $\alpha_{NB}$  of the Petri Net transitions. In particular for a negative influence, the transition  $\vec{x}' = \vec{x} - t_i$  in the integer vector domain, is abstracted in the boolean vector domain by the zero/non zero abstraction with the two transitions  $\vec{x}' = \vec{x}$  or  $\vec{x}' = \vec{x} - t_i$ .  $\square$

**Corollary 1.** *If a behaviour is not possible in the positive boolean semantics of an influence system, it is also not possible in the Petri Net and stochastic semantics for any forces.*

### 3.3 Expressive Power Compared to Reaction Systems

**Theorem 3.** *Any (well-formed) influence system with forces can be represented by a (well-formed) reaction system, with the same boolean, Petri Net, stochastic and differential semantics.*

*Proof.* Let us represent a positive influence  $f$  for  $P/I \rightarrow t$  by a catalytic synthesis reaction  $f$  for  $P/I \Rightarrow P + t$ .

Similarly, let us represent a negative influence  $f$  for  $P/I \rightarrow t$ , by an active degradation reaction  $f$  for  $P + t/I \Rightarrow P$ .

It is straightforward to verify that the boolean, Petri Net, stochastic as well as differential semantics recalled and defined above are the same.

Furthermore, the well-formedness condition is preserved. Indeed, this property only depends on the forces/rate functions and on the reactants/inhibitors, which do not change through that transformation thanks to the condition that in *well-formed* influence systems. In addition,  $t \in P$  in the case of the negative influence.  $\square$

This theorem shows that an influence system can be simulated by a reaction system in the hierarchy of semantics. The converse does not hold for the discrete semantics for two reasons. First, self-loop transitions cannot be represented in the Petri Net semantics of influence systems (see remark above made with the definition). Second, in the discrete (stochastic, Petri Net, boolean) semantics of the reaction  $C + D \Rightarrow A + B$  (the rate is omitted), we have a transition from the state  $(A, B, C, D) = (0, 0, 1, 1)$  to  $(1, 1, 0, 0)$  which is obviously not possible in an influence system since only one variable can change in one transition.

Let us call a reaction  $(R, I, P, f)$  *unitary* if the support of  $P - R$  is a singleton with a result equal to either 1 or -1.

**Proposition 1.** *Any (well-formed) system of unitary reactions can be represented by a (well-formed) influence system with the same boolean, Petri Net, and stochastic semantics.*

*Proof.* Let us associate with each unitary reaction  $(R, I, P, f)$ , with  $P(t) - R(t) = \pm 1$  and  $P(x) = R(x)$  for all  $x \neq t$ , the influence  $(R, I, t, \sigma, f)$  where  $\sigma$  is the sign of  $P(t) - R(t)$ .

The stochastic and Petri Net transitions associated with the unitary reaction  $(R, I, P, f)$  apply on discrete states satisfying  $\vec{x} \geq R$  and  $\vec{x}' = \vec{x} - R + P = \vec{x} \sigma t$ , i.e. the same transitions as for the influence  $(R, I, t, \sigma, f)$ .

The boolean transitions associated with the unitary reaction  $(R, I, P, f)$  apply on discrete states satisfying  $\vec{x} \geq R$  and  $\vec{x}' = \vec{x} - S + P$  for all  $S \leq R$ .

If  $P(t) - R(t) = +1$  then  $\vec{x}' = \vec{x} + t$ .

If  $P(t) - R(t) = -1$  then either  $\vec{x}' = \vec{x}$  or  $\vec{x}' = \vec{x} - t$ .

Therefore in both cases, the boolean transitions associated with the reaction  $(R, I, P, f)$  are the same as those associated with the influence  $(R, I, t, \sigma, f)$ .  $\square$

Interestingly, the converse holds for the differential semantics with all generality:

**Theorem 4.** *Under the differential semantics, (well-formed) influence and reaction systems have the same expressive power.*

*Proof.* For each reaction  $(R, I, P, f)$  of a given reaction system, let us add the following influences:

$$\begin{aligned} &(R, I, x_i, +, (P(i) - R(i)) \times f) \text{ when } P(i) - R(i) > 0 \\ &(R, I, x_i, -, (R(i) - P(i)) \times f) \text{ when } P(i) - R(i) < 0 \end{aligned}$$

The associated ODE system collects all  $(P_i - R_i) \times f$  exactly as in the differential semantics of the original reaction system.

Furthermore, it is easy to check that these influences are well-formed since the original reaction is well-formed and the force is only a positive integer multiplied by the original rate function.  $\square$

This theorem shows that as far as the differential semantics is concerned, the influence systems have the same expressive power as reaction systems and there is no theoretical reason to develop a reaction model. This does not mean that there is a canonical reaction system associated with an influence system. Generally, different implementations with reactions are possible without changing the differential semantics. They represent extra information which are irrelevant to the analysis or simulation of the differential equations, but which provide richer discrete dynamics.

### 3.4 Semantics with Negation

Contrarily to reaction systems, it is usual in influence systems and regulatory networks *à la* Thomas, to adopt a strict interpretation of the negative sources of an influence as tests for absence in the enabling condition, and of the negative influences as negation of the target in the boolean semantics. These choices are made at the expense of losing the abstraction relationship from the Petri Net semantics to the boolean semantics (Thm. 2) since for instance, a positive influence like  $A/2B \rightarrow C$  can activate  $C$  with  $B$  present at level 1 in the Petri Net semantics, but not in the boolean semantics which is thus no longer an abstraction of the Petri Net semantics.

Formally, the *boolean with negation semantics* of an influence system is defined by the following transition system:

$$\begin{aligned} &\forall (P_i, I_i, t_i, \sigma_i, f_i), \vec{x} \longrightarrow_{BN} \vec{x}' \\ &\text{if } \vec{x} \geq P_i, \vec{x} <_0 I_i, \vec{x}' = \vec{x} \sigma_i t_i \end{aligned}$$

The stochastic and Petri Net semantics with negation are defined similarly by adding the condition  $\vec{x} <_0 I_i$  on the negative sources to the transition.

It is worth remarking that this strict interpretation of the negative sources makes it possible to express influence systems with *bounded multilevel semantics*, as considered in tools like GINsim [9] for instance, just by adding a negative source  $k * x$  to all the positive influences on  $x$ , with some a



level coefficient bound  $k$ . This is not possible in the positive semantics of the previous section.

That interpretation of negative sources by negation increases the expressive power of influence systems. Let us call a *unitary transition system*, a transition system that updates at most one variable of  $\vec{x}$  in each transition<sup>5</sup>.

**Proposition 2.** *Any unitary boolean transition system can be represented by an influence system under the boolean semantics with negation.*

*Proof.* It is sufficient to notice that since a unitary boolean transition  $s \rightarrow_{BN} s'$  changes at most one species, say  $t_i$ , from  $s$  to  $s'$ , it can be represented by

- either a positive influence,  $(P, I, t_i, +)$ , if  $s'(t_i) = 1$ ,
- or a negative influence,  $(P, I, t_i, -)$ , if  $s'(t_i) = 0$ ,

with  $P = \{x \mid s(x) = 1\}$  and  $I = \{x \mid s(x) = 0\}$ .  $\square$

### 3.5 Functional Semantics à la Thomas

The boolean and multilevel semantics defined by René Thomas originally for gene networks [6], is *functional*, in the sense that the next state  $\vec{x}'$  is defined by a function  $\phi(\vec{x})$  of the previous state, not a relation. In this setting, the synchronous semantics is deterministic, and the non-deterministic asynchronous semantics is obtained by interleaving, by considering all the possible transitions that change the value of *exactly one* of the genes at a time. A truly non-deterministic influence system such as  $\{(A, \emptyset, B, +, f), (A, \emptyset, B, -, g)\}$ , for which the transition relation is not a function, cannot be represented. Thomas's setting excludes self-loops in the state transition graph and all steady states are stable (i.e. terminal states).

**Proposition 3.** *The set of boolean transition systems definable by Thomas's regulatory networks is the set of unitary boolean transition systems without self-loops.*

*Proof.* A Thomas's transition graph is necessarily unitary and without self-loops since each transition changes the boolean value of exactly (not at most) one variable at a time.

The converse follows from Prop. 2 by excluding the possibility of having self-loop transitions which change no variable.  $\square$

This restriction to transition functions is even more striking in Thomas's multilevel setting, where the above system can (in the discrete semantics) have transitions from  $(1, 1)$  both to  $(1, 0)$  and to  $(1, 2)$ . That would necessitate the corresponding logical parameter for  $B$  to be at the same time  $< 1$  and  $> 1$ . It is worth noting that despite this subtle restriction in the expressive power, the logical formalism of Thomas is successfully used in a wide variety of models [22], [29], [30], [31] in systems biology.

### 3.6 Influence Graph of an Influence System

The differential influence graph of an influence system is defined though the differential semantics, as in Def. 3. We get a similar equivalence result with the following syntactical definition of the influence graph:

5. It is worth remarking that in a unitary transition system, the state transition graph lives on a hypercube (see e.g. Fig. 4 in Sec. 5 below).

**Definition 7.** *The syntactical influence graph (SIG) associated with an influence system  $\mathcal{I}$  is the graph having for vertices the molecular species, and for edges the following set of positive and negative influences:*

$$\begin{aligned} \{A \rightarrow^+ B \mid & \exists (P_i, I_i, B, \sigma_i, f_i) \in \mathcal{I} \\ & (A \in P_i \text{ and } \sigma_i = +) \\ & \text{or } (A \in I_i \text{ and } \sigma_i = -)\} \\ \cup \{A \rightarrow^- B \mid & \exists (P_i, I_i, B, \sigma_i, f_i) \in \mathcal{I} \\ & (A \in P_i \text{ and } \sigma_i = -) \\ & \text{or } (A \in I_i \text{ and } \sigma_i = +)\} \end{aligned}$$

**Proposition 4.** *For a well-formed influence system such that the syntactical influence graph contains no conflict, the syntactical and differential influence graphs are identical.*

*Proof.* We have

$$\begin{aligned} \dot{x}_B &= \sum_{(P_i, I_i, x_B, +, f_i) \in \mathcal{I}} f_i - \sum_{(P_j, I_j, x_B, -, f_j) \in \mathcal{I}} f_j \\ \frac{\partial \dot{x}_B}{\partial x_A} &= \sum_{(P_i, I_i, x_B, +, f_i) \in \mathcal{I}} \frac{\partial f_i}{\partial x_A} - \sum_{(P_j, I_j, x_B, -, f_j) \in \mathcal{I}} \frac{\partial f_j}{\partial x_A} \end{aligned}$$

Since the SIG does not have any conflict,  $A \rightarrow^+ B$  is in the SIG if and only if

$$\begin{aligned} \frac{\partial \dot{x}_B}{\partial x_A} &= \sum_{(P_i, I_i, x_B, +, f_i) \in \mathcal{I}, A \in P_i, A \notin I_i} \frac{\partial f_i}{\partial x_A} \\ &\quad - \sum_{(P_j, I_j, x_B, -, f_j) \in \mathcal{I}, A \notin P_j, A \in I_j} \frac{\partial f_j}{\partial x_A} \end{aligned}$$

and a similar reasoning can be made for  $A \rightarrow^- B$ .

Now, since the influence system is well-formed, all terms of the left-hand sum are non-negative ( $A \notin I_i$ ) and strictly positive for some points  $\vec{x}_i$ .

Similarly, all terms of the right-hand sum are non-positive ( $A \notin P_j$ ) and strictly negative for some  $\vec{x}_j$ .

We have that  $A \rightarrow^+ B$  in the SIG iff the above sum has at least one term, that is equivalent to the existence of some  $\vec{x}$  in the state space where one of the terms above is non-null, in which case  $\frac{\partial \dot{x}_B}{\partial x_A} > 0$ , i.e.,  $A \rightarrow^+ B$  is in the DIG.  $\square$

One can also easily check that in the representation of an influence system by a reaction system (Thm. 3), the influence graphs of both systems are the same.

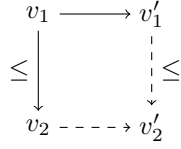
## 4 PROPERTIES OF THE POSITIVE SEMANTICS AND COMPUTATION OF BOOLEAN ATTRACTORS

In this section, we focus on the positive boolean semantics of influence systems and study its properties. In particular we show that the boolean attractors of the positive semantics contain the attractors of the functional semantics with negation à la Thomas and provide an efficient pruning algorithm for computing them. Let us recall that  $\leq$  denotes the point-wise order on  $\{0, 1\}$  coordinates of vectors representing states.

**Proposition 5 (Monotonicity).** *The positive boolean semantics of influence systems is monotonic: let  $I$  be an influence system over  $S = \{x_1, \dots, x_s\}$  and  $v_1, v_2$  be two boolean states, i.e., vectors of  $\mathbb{B}^s$*

$$\text{if } v_1 \leq v_2 \text{ then } \forall v'_1, v_1 \rightarrow v'_1, \exists v'_2, v'_1 \leq v'_2 \text{ and } v_2 \rightarrow v'_2$$





*Proof.* One can simply notice that since there are no negations in the enabling conditions, any influence that is enabled in  $v_1$  is also enabled in  $v_2$ .  $\square$

It is worth noticing that this monotonicity property for transitions is fundamentally different from that of monotone dynamical systems [32] which are deterministic, and therefore impose the monotonicity property on the *unique* image of  $v_1$  and  $v_2$ . In our setting, Prop. 5 states that there exists some  $v'_2 \geq v'_1$ , but the existence of negative influences in the system permits that some other images of  $v_2$  might not be greater than  $v'_1$ . Nevertheless, we have

**Proposition 6** (Greatest element). *Let  $C$  be a Terminal Strongly Connected Component (TSCC) of the state transition graph of a positive influence system, then  $C$  has a greatest element:  $\exists v_0 \in C, \forall v \in C, v \leq v_0$*

*Proof.* Let us prove this proposition by contradiction: assume that there are two incomparable maximal elements  $v_1$  and  $v_2$  in  $C$ .

Since  $C$  is strongly connected there is a path from  $v_1$  to  $v_2$  and along that path a state  $v_3$  and its successor in the path  $v_4$  such that  $v_3 \leq v_1$  and  $v_4 \not\leq v_1$ , as  $v_2 \not\leq v_1$ .

Now, using Prop. 5 we get that  $v_1 \rightarrow v'_1$  with  $v_4 \leq v'_1$  and  $v'_1 \in C$  since  $C$  is terminal.

However,  $v'_1$  is either greater or less than  $v_1$  since it is the result of applying a single influence.

If  $v_1 < v'_1$  we have a contradiction as we supposed  $v_1$  maximal.

If  $v'_1 \leq v_1$  we get  $v_4 \leq v_1$  by transitivity and that is also contradictory.  $\square$

**Corollary 2.** *To enumerate the attractors, i.e., TSCCs, of a positive influence system, it is enough to check the strongly connected components (SCCs) of states that have no strictly increasing transition.*

*Proof.* This is an immediate consequence of Prop. 6 as each TSCC can be represented by its greatest element which has no strictly increasing transition.  $\square$

Notice that stable states are a particular case with no strictly decreasing transition either. Moreover, any strictly decreasing transition should be “reversible” for the SCC to be a TSCC. This allows us to rule out potential TSCC candidates without exploring their whole SCC in Alg. 1 and Alg. 2 (implementation available in Biocham v4).

**Proposition 7.** *Given an influence system, there is at least one TSCC of the original influence system in each TSCC of its positive semantics.*

*Proof.* The positive semantics only adds transitions by enabling more influences, it can therefore only merge TSCCs.  $\square$

This result suggests finding complex attractors of non-positive systems, such as logical models *à la* Thomas [7], [8],

**Algorithm 1** TSCC maximal elements candidates enumeration algorithm

---

```

procedure LIST_TSCC_CANDIDATES
   $Constraints \leftarrow \{P \wedge \neg I \implies t \mid (P, I, +, t, f) \in \mathcal{I}\}$ 
   $\triangleright$  Enabled positive influences must not change the state
   $Candidates \leftarrow \text{ENUMERATE\_SOLUTIONS}(Constraints)$ 
  for  $C \in Candidates$  do
    if  $C$  has no strictly decreasing transition then
       $C$  is a stable steady state
    else if  $C$  has a non-reversible strictly decreasing transition then
       $C$  is not in a TSCC
    else
       $C$ 's SCC must be explored to check if it is a TSCC
    end if
  end for
end procedure

function ENUMERATE_SOLUTIONS( $Constraints$ )
  Iteratively solve by SAT/CP the Constraint Satisfaction Problem (CSP) defined by  $Constraints$ 
  return The set of solutions
end function

```

---

by enumerating the greatest elements of the TSCCs of their positive boolean semantics, and then looking for attractors of the original system. This approach provides an over-approximation of the attractors and is complementary to recent works which provide lower-bounds on their number [33].

One might note however that the reversibility constraint on decreasing transitions can easily be stated as a constraint of our Constraint Satisfaction Problem (CSP). Now, checking that an SCC is terminal can be done through model checking, and verifying if there is a singular TSCC inside a positive TSCC is also a reachability problem. Putting all this together we reach a refined version, Alg. 2 of our original algorithm, Alg. 1 that now guarantees that an SCC it finds is terminal, and also checks for inclusion of stable states. Alg. 2 therefore provides a lower bound on the number of complex attractors, and an over-approximation of those.

## 5 EXAMPLES AND EVALUATION

### 5.1 Influence Model of the p53/Mdm2 DNA Damage Repair System [19]

The p53/Mdm2 DNA damage repair system is an interesting oscillatory system which has been modeled by differential equations by Ciliberto et al. in [18]. Fig. 5.1 shows the diagram of interaction which has been transcribed in a reaction model. Note that this model contains explicit transport reactions of Mdm2 between the cytoplasm and the nucleus, and that p53 is an inhibitor of the phosphorylation reaction of Mdm2 in the cytoplasm (reaction of the form  $k * \text{Mdm2} / (j + p53)$  for  $\text{Mdm2} / p53 \Rightarrow \text{Mdm2p}$ ). Fig. 5.1 depicts the influence graphs (inferred by Def. 4) of this reaction model, and of a model reduction obtained by

- by merging p53 and p53u in P,
- deleting p53uu,

**Algorithm 2** Refined TSCC search algorithm

---

```

procedure LIST_TSCC_CANDIDATES
  Stable  $\leftarrow$  ENUMERATESTABLE()
  Constraints  $\leftarrow \{P \wedge \neg I \implies t \mid (P, I, +, t, f) \in \mathcal{I}\}$ 
   $\triangleright$  Enabled positive influences must not change the state
  Constraints  $\leftarrow$  Constraints  $\cup \{P \wedge \neg I \implies t \vee$ 
 $\vee_i P_i \wedge \neg I_i \mid (P, I, -, t, f), (P_i, I_i, +, t, f_i) \in \mathcal{I}, t \notin P_i\}$ 
   $\triangleright$  Enabled negative influences must be reversible
  Candidates  $\leftarrow$  ENUMERATESOLUTIONS(Constraints)
  for C  $\in$  Candidates do
    if C  $\in$  Stable then
      C is a stable steady state
    else if C  $\not\models AG(EF(C))$  (positive semantics) then
      C is not in a TSCC
    else
      C is in a positive TSCC
      if  $\forall S \in \text{Stable}, C \models \neg EF(S)$  (negative semantics) then
        there is at least one negative non-trivial
        TSCC (a complex attractor) in C's TSCC
      end if
    end if
  end for
end procedure

function ENUMERATESTABLE
  Iteratively solve by SAT/CP the CSP corresponding to
  the absence of state-changing enabled influence
  return The set of solutions
end function

function ENUMERATESOLUTIONS(Constraints)
  Iteratively solve by SAT/CP the CSP defined by
  Constraints
  return The set of solutions
end function

```

---

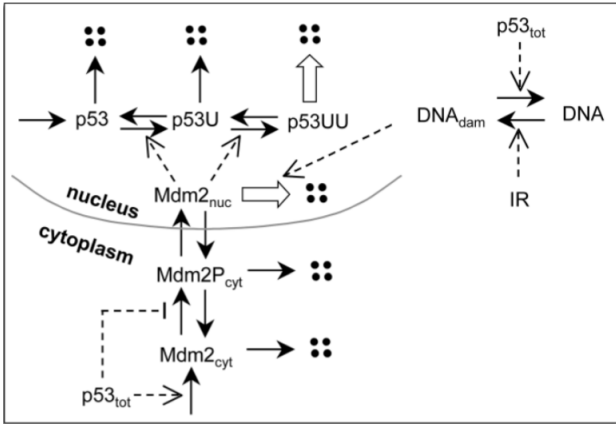


Fig. 1. Diagram of the p53/Mdm2 DNA damage repair system [18].

- merging Mdm2pc and Mdm2n in N,
- and renaming Mdm2c in C and DNAdam in D,

i.e. a subgraph epimorphism (SEPI) reduction [34].

In [19] and [20] two simplified influence models were manually derived from this model, and analyzed with boolean, multilevel and differential semantics. We illustrate here the search for TSCCs with two versions corresponding to the boolean and multilevel semantics of the influence

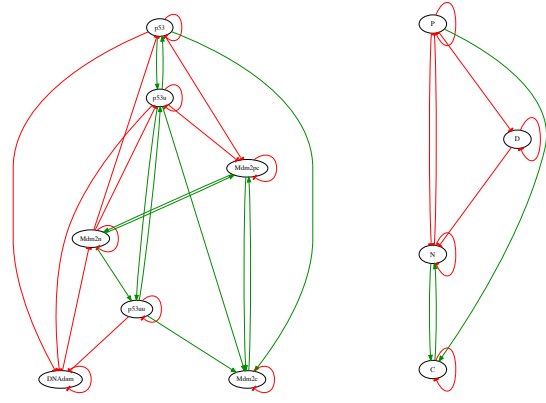


Fig. 2. Influence graphs (Def. 4) of the reaction model of Ciliberto et al. [18], and of a SEPI reduction of this reaction model [34].

model of [19]. In the first model, we simply transcribe the graph of Fig. 4 of the authors as a boolean influence system. We therefore ignore the multi-level aspect they developed. In the second model, we add some activations on p53 and DNA-damage, and an inhibition on cytoplasmic Mdm2, in order to take into account some basal state of the model.

Fig. 3 depicts the influence graph of this models. Interestingly, we can see that it matches nearly exactly the influence graph of the reduced model of Ciliberto et al. The only differences are the absence of the negative self-loops, and of the positive influence from N to C (corresponding to the transport of Mdm2 from the nucleus to the cytoplasm) which were neglected.

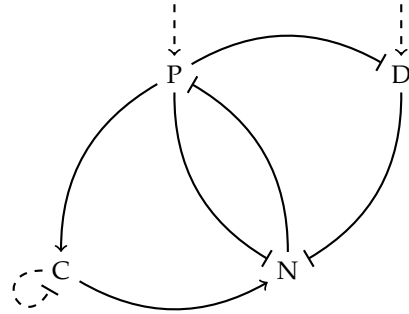


Fig. 3. Influence graph displayed in Fig. 4 of [19], without the activation multi-levels. The dashed influences are those added in our second version of the model.

The following Biocham v4 trace shows the TSCCs computed by Alg. 2 in both influence systems. The CPU times are given in seconds and have been obtained on a 2010 Max-mini, as all the CPU times given in this paper.

```

biocham: load_biocham(kaufman) .
biocham: list_influences.
[0] P -> C
[1] C -> N
[2] N -< P
[3] P -< N
[4] P -< D

```

```

[5] D -< N
biocham: list_tscc_candidates.
[C-0,D-0,N-0,P-0] stable
[C-0,D-0,N-1,P-0] stable
[C-0,D-1,N-0,P-0] stable
[C-1,D-0,N-1,P-0] stable
[C-1,D-1,N-1,P-0] terminal (positive)
contains a complex attractor
Time: 0.107
biocham: add_biocham(kaufman2).
biocham: list_influences.
[0] P -> C
[1] C -> N
[2] N -< P
[3] P -< N
[4] P -< D
[5] D -< N
[6] _ -> P
[7] _ -> D
[8] C -< C
biocham: list_tscc_candidates.
[C-1,D-1,N-1,P-1] terminal (positive)
contains a complex attractor
Time: 0.048

```

Our algorithm shows that there is in each case a single complex attractor accordingly to [19], and four stable steady states in the first case. In [20], this influence model was further extended with differential and stochastic dynamics which could be directly represented by influence forces in our setting.

## 5.2 Influence Model of the Mammalian Circadian Clock [21]

A nice example of the use of logical models *à la Thomas* is the recent paper by Comet et al. [21] studying different variants of small models of the circadian rhythms in mammals. Starting from the following 4-variable ODE model, which is a very simplified version of the 16-variable model from [35], they obtain a small Thomas influence network.

$$\begin{aligned}
\frac{dP_C}{dt} &= \frac{K^n}{K^n + PC_N^n} v_1 - k_3 P_C C_C + k_4 P_C C - k_{d1} P_C \\
\frac{dC_C}{dt} &= \frac{K^n}{K^n + PC_N^n} v_2 - k_3 P_C C_C + k_4 P_C C - k_{d2} C_C \\
\frac{dP_C C}{dt} &= k_3 P_C C_C - k_4 P_C C - k_1 P_C C + k_2 P_C N - k_{d3} P_C C \\
\frac{dP_C N}{dt} &= k_1 P_C C - k_2 P_C N - k_{d4} P_C N
\end{aligned}$$

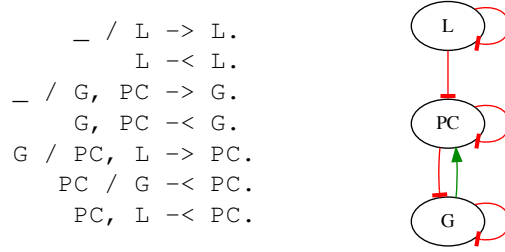
In the obtained logical model they introduce light and further simplify the Per and Cry proteins in order to study light entrainment:

“Remarkably, from the pedagogical point of view, this very simple model provides a very pertinent causal explanation of the robustness of the circadian cycle to variations of day length. This model allows to stand back, and so, to better understand the other models.”

A final refinement introduces delays in order to constrain the light-dark cycles (i.e. the reversible arrows from  $L = 0$

to  $L = 1$ ), but this hybrid part is out of the scope of this paper.

A direct import in Biocham v4 of the logical model without delays (cf. Section 5 of [21]) gives the following influence system with negative sources, shown below with the corresponding influence graph:



The positive semantics of this system is close to the original boolean semantics with negation *à la Thomas* of the model. Only a few state transitions become reversible in the positive boolean semantics, while they are irreversible in the original boolean semantics with negation *à la Thomas* of the model, as depicted in Fig. 4.

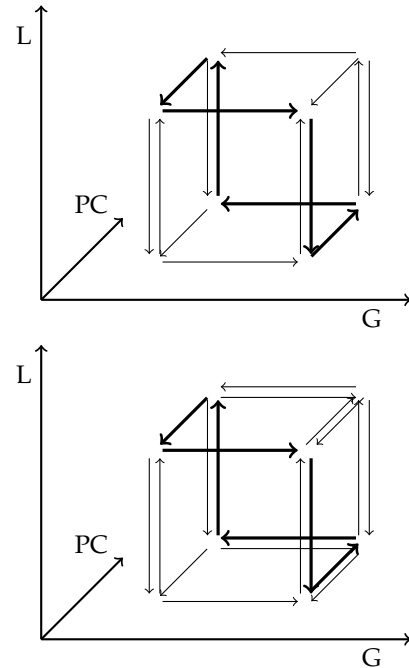


Fig. 4. State transition graphs of the model under the boolean semantics with negation *à la Thomas*, similar to Fig. 7 of [21] (top), and the positive boolean semantics, where some state transitions have become reversible (bottom). The bold circuit represents the classical 24h day-night cycle that is obtained as standard cycle when delays are introduced. The other transitions correspond to recovery pathways in a jet lag context for instance.

They both have a single TSCC: the whole state transition graph represented by the vector  $(1, 1, 1)$  that is found as sole candidate:

```

biocham: load_biocham(bernot).
biocham: list_tscc_candidates.
[G-1,L-1,PC-1] terminal (positive)

```

contains a complex attractor  
Time: 0.113

The approximation introduced by the positive boolean semantics can be explained by quantitative dynamics considerations. For instance, when  $G$  is on, the transcription leading to the PER-CRY complexes is stimulated, however [21] explains that these complexes can only migrate to the nucleus in absence of light. This *absence* cannot be checked in a positive semantics model, however the consensus mechanistic process is rather thought to be a modulation of PER transcription by light (see for instance [35] for the mammalian case). Being purely quantitative, it is not easy to take into account such a regulation in a boolean model except with the reversible activation of  $PC$  when  $G$  is on, whether  $L$  is on or not. This is what happens in our positive model as can be seen in the right panel of Fig. 4, and it is similar to what happens for the light in the original model.

The same reasoning explains the reversible inactivation of  $G$  when  $PC$  is active. Indeed there is a basal synthesis of  $G$  that cannot check, in a positive setting, that  $PC$  is inactive in order to activate the genes. Once again, the mechanistic process is a quantitative inhibition of the CLOCK-BMAL1 complexes by PER-CRY and a conservative boolean approximation of that process is reflected by the reversible activation of  $G$  in presence of  $PC$ .

In [21], the authors also restrict the possible behaviours by introducing delays for the boolean transitions, which could be considered as a further expansion of the formalism.

### 5.3 Influence Model of MAPK Signaling Network [22]

As a more challenging example, we present here the application of our methods on a logical model of the Mitogen Activated Protein Kinase (MAPK) network presented in [22]. Starting, once again, from a “comprehensive reaction map” of 232 species, the authors design a reduced logical model in Thomas’ framework. This 53-variable model, also used as benchmark in [33], describes the influence of the MAPK cascade (and of its inputs, like EGFR over-expression and FGFR3 activating mutation) on cell-fate decision in urinary bladder cancer.

Alg. 2 gives the following results, with a total runtime of about 106 seconds:

- 12 steady-states are found. These are the most important features of the model for its original publication (cell-fate) and are easy to find whatever the method (MDDs in GINsim, ASP in [33] and SAT or CSP for us).
- there are 21 states without any strictly increasing outgoing transition found by constraint solving;
- of those 15 are in **terminal** SCCs as proven by a single call to a model checker;
- and of those TSCCs, only 9 contain one of the previously found steady-states, leading to at least 6 complex attractors.

The result is therefore the same bound as the one found in [33]. However, proving that bound exact would be much slower in our case than in [36], since our approximation of the negative SCCs by positive-SCCs is not as well suited to model checking as the trap-spaces which can be directly encoded as BDDs.

## 6 CONCLUSION

In this paper, we hope to have clarified some differences between influence systems and reaction systems, and especially some subtle discrepancies between the precise boolean semantics that have been considered in the literature. As far as the modeling of one biological system is concerned, a modeler can work with one formalism and one tool to investigate his/her questions about the system. Nevertheless, as soon as different modeling tools are to be used, or the model has to be communicated and reused or coupled to another one for another purpose, understanding and mastering these discrepancies in the semantics of the interactions becomes crucial.

We have shown that, for influence systems and reaction systems with inhibitors, one can obtain a hierarchy of semantics which goes from the concrete stochastic semantics to a Petri net, and then a positive boolean semantics in which the inhibitors of the reactions or influences are just ignored. This is consistent with the fact that the inhibitors decrease the rate or force in the quantitative semantics, but do not really prevent the reaction or influence from proceeding. This convention thus ensures that all discrete behaviours are approximated when we go up in the abstractions of the hierarchy of semantics, and that if a behaviour is not possible in the positive boolean semantics (which can be checked by model checking methods for instance) it is not possible in the stochastic semantics for any forces. Furthermore, we have shown that in the positive boolean semantics, the monotonicity of the transition relation allows us to enumerate complex attractors efficiently by restricting the search to the greatest elements candidates.

On the other hand, the boolean semantics with negation lead to a more expressive formalism in which any unitary boolean transition system can be encoded, and for which a similar hierarchy of discrete semantics with negation can also be developed. The stochastic and Petri Net semantics with negation interpret inhibitors as negative conditions, but with no consistent link to the differential semantics.

Furthermore, the boolean semantics with negation *à la* Thomas, contains a restriction on the definition of the transition relation by a function, as opposed to a relation, which limits the sources of non-determinism. With the functional restriction, we have proven that each terminal strongly connected component of the transition graph (TSCC) in the positive semantics contains at least one TSCC of the semantics *à la* Thomas, and thus that our TSCC computation algorithm can be used to prune the search space in this setting also.

We have also shown that reaction systems and influence systems have the same expressive power under the differential semantics. This means that, as far as the differential equations are concerned, the details given in the reactant-product structure of a reaction system are not necessary, and that the same differential equations can be derived from an influence system with forces.

## ACKNOWLEDGMENTS

We are grateful to Paul Ruet and Denis Thieffry for interesting discussions on Thomas’s framework, and to the reviewers for their comments. This work was partially

supported by ANR project Hyclock under contract ANR-14-CE09-0011, and PASPA-DGAPA-UNAM, Conacyt grants 221341 and 261225.

## REFERENCES

- [1] M. Hucka *et al.*, "The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models," *Bioinformatics*, vol. 19, no. 4, pp. 524–531, 2003. [Online]. Available: <http://sbml.org/>
- [2] N. le Novère, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M. Schilstra, B. Shapiro, J. L. Snoep, and M. Hucka, "BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems," *Nucleic Acid Research*, vol. 1, no. 34, pp. D689–D691, Jan. 2006.
- [3] F. Fages and S. Soliman, "Abstract interpretation and types for systems biology," *Theoretical Computer Science*, vol. 403, no. 1, pp. 52–70, 2008. [Online]. Available: <http://lifeware.inria.fr/~fages/Papers/FS07tcs.pdf>
- [4] L. Glass and S. A. Kauffman, "The logical analysis of continuous, non-linear biochemical control networks," *Journal of theoretical Biology*, vol. 39, no. 1, pp. 103–129, 1973.
- [5] R. Thomas, "Boolean formalisation of genetic control circuits," *Journal of Theoretical Biology*, vol. 42, pp. 565–583, 1973.
- [6] R. Thomas and R. D'Ari, *Biological Feedback*. CRC Press, 1990.
- [7] E. Remy, P. Ruet, and D. Thieffry, "Graphic requirements for multi-stability and attractive cycles in a boolean dynamical framework," *Advances in Applied Mathematics*, vol. 41, no. 3, pp. 335–350, 2008.
- [8] P. Ruet, "Local cycles and dynamical properties of boolean networks," *Mathematical Foundations of Computer Science*, vol. 26, no. 4, pp. 702–718, 2016.
- [9] A. Naldi, D. Berenguier, A. Fauré, F. Lopez, D. Thieffry, and C. Chaouiya, "Logical modelling of regulatory networks with GINsim 2.3," *Biosystems*, vol. 97, no. 2, pp. 134–139, 2009.
- [10] G. Batt, B. Besson, P. Ciron, H. de Jong, E. Dumas, J. Geiselmann, R. Monte, P. Monteiro, M. Page, F. Rechenmann, and D. Ropers, "Genetic Network Analyzer: a tool for the qualitative modeling and simulation of bacterial regulatory networks," in *Bacterial Molecular Networks*. Springer, 2012, pp. 439–462.
- [11] D. A. Rosenblueth, S. Muñoz, M. Carrillo, and E. Azpeitia, "Inference of Boolean networks from gene interaction graphs using a SAT solver," in *AlCoB 2014: Proceedings of the 1st International Conference on Algorithms for Computational Biology*, ser. Lecture Notes in Bioinformatics, vol. 8542. Springer-Verlag, 2014, pp. 235–246.
- [12] C. Chaouiya, "Petri net modelling of biological networks," *Briefings in Bioinformatics*, vol. 8, no. 210, 2007.
- [13] B. Chazelle, "Natural algorithms and influence systems," *Communications of the ACM*, vol. 55, no. 12, pp. 101–110, Dec. 2012.
- [14] A. Naldi, R. Remy, D. Thieffry, and C. Chaouiya, "A reduction method for logical regulatory graphs preserving essential dynamical properties," in *CMSB'09: Proceedings of the seventh international conference on Computational Methods in Systems Biology*, ser. Lecture Notes in Bioinformatics, vol. 5688. Springer-Verlag, 2009, pp. 266–280.
- [15] N. le Novère, M. Hucka, H. Mi, S. Moodie, F. Schreiber, A. Sorokin, E. Demir, K. Wegner, M. I. Aladjem, S. M. Wimalaratne, F. T. Bergman, R. Gauges, P. Ghazal, H. Kawaji, L. Li, Y. Matsuoka, A. Villegier, S. E. Boyd, L. Calzone, M. Courtot, U. Dogrusoz, T. C. Freeman, A. Funahashi, S. Ghosh, A. Jouraku, S. Kim, F. Kolpakov, A. Luna, S. Sahle, E. Schmidt, S. Watterson, G. Wu, I. Goryanin, D. B. Kell, C. Sander, H. Sauro, J. L. Snoep, K. Kohn, and H. Kitano, "The systems biology graphical notation," *Nature Biotechnology*, vol. 27, no. 8, pp. 735–741, Aug. 2009.
- [16] F. Fages and S. Soliman, "From reaction models to influence graphs and back: a theorem," in *Proceedings of Formal Methods in Systems Biology FMSB'08*, ser. Lecture Notes in Computer Science, no. 5054. Springer-Verlag, Feb. 2008. [Online]. Available: <http://lifeware.inria.fr/~fages/Papers/FS08fmsb.pdf>
- [17] P. Cousot and R. Cousot, "Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints," in *POPL'77: Proceedings of the 6th ACM Symposium on Principles of Programming Languages*. New York: ACM Press, 1977, pp. 238–252, Los Angeles.
- [18] A. Ciliberto, F. Capuani, and J. J. Tyson, "Modeling networks of coupled enzymatic reactions using the total quasi-steady state approximation," *PLOS Computational Biology*, vol. 3, no. 3, p. e45, Mar. 2007.
- [19] W. Abou-Jaoudé, D. A. Ouattara, and M. Kaufman, "From structure to dynamics: frequency tuning in the p53-Mdm2 network: I. logical approach," *Journal of theoretical biology*, vol. 258, no. 4, pp. 561–577, 2009.
- [20] —, "From structure to dynamics: frequency tuning in the p53-Mdm2 network: II differential and stochastic approaches," *Journal of theoretical biology*, vol. 264, pp. 1177–1189, 2010.
- [21] J.-P. Comet, G. Bernot, A. Das, F. Diener, C. Massot, and A. Cessieux, "Simplified models for the mammalian circadian clock," *Procedia Computer Science*, vol. 11, pp. 127–138, 2012.
- [22] L. Grieco, L. Calzone, I. Bernard-Pierrot, F. Radvanyi, B. Kahn-Perlès, and D. Thieffry, "Integrative modelling of the influence of mapk network on cancer cell fate decision," *PLOS Computational Biology*, vol. 9, no. 10, p. e1003286, Oct. 2013.
- [23] F. Fages, S. Gay, and S. Soliman, "Inferring reaction systems from ordinary differential equations," *Theoretical Computer Science*, vol. 599, pp. 64–78, Sep. 2015. [Online]. Available: <http://lifeware.inria.fr/~fages/Papers/FGS14tcs.pdf>
- [24] M. Katsumata, "Graphic representation of Botts-Morales equation for enzyme-substrate-modifier system," *Journal of Theoretical Biology*, vol. 36, no. 2, pp. 327–338, 1972.
- [25] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *Journal of Physical Chemistry*, vol. 81, no. 25, pp. 2340–2361, 1977.
- [26] N. Chabrier and F. Fages, "Symbolic model checking of biochemical networks," in *CMSB'03: Proceedings of the first workshop on Computational Methods in Systems Biology*, ser. Lecture Notes in Computer Science, C. Priami, Ed., vol. 2602. Rovereto, Italy: Springer-Verlag, Mar. 2003, pp. 149–162. [Online]. Available: <http://lifeware.inria.fr/~fages/Papers/CFG02cmsb.pdf>
- [27] G. Bernot, J.-P. Comet, and Z. Khalis, "Gene regulatory networks with multiplexes," in *Proc of European Simulation and Modelling Conference, ESM'2008*, 2008, pp. 423–432.
- [28] F. Fages, T. Martinez, D. Rosenblueth, and S. Soliman, "Influence systems vs reaction systems," in *CMSB'16: Proceedings of the fourteenth international conference on Computational Methods in Systems Biology*, ser. Lecture Notes in Bioinformatics, N. P. E. Bartocci, P. Lio, Ed., vol. 9859. Springer-Verlag, Sep. 2016, pp. 98–115. [Online]. Available: <https://hal.inria.fr/hal-01378470>
- [29] P. Traynard, A. Fauré, F. Fages, and D. Thieffry, "Logical model specification aided by model-checking techniques: application to the mammalian cell cycle regulation," *Bioinformatics*, vol. 32, no. 17, pp. i772–i780, 2016. [Online]. Available: <https://hal.inria.fr/hal-01378465>
- [30] A. Naldi, J. Carneiro, C. Chaouiya, and D. Thieffry, "Diversity and plasticity of th cell types predicted from regulatory network modelling," *PLoS Computational Biology*, vol. 6, no. 9, p. e1000912, Sep. 2010.
- [31] A.-G. González, C. Chaouiya, and D. Thieffry, "Qualitative dynamical modelling of the formation of the anterior-posterior compartment boundary in the drosophila wing imaginal disc," *Bioinformatics*, vol. 24, pp. 234–240, 2008.
- [32] D. Angeli and E. D. Sontag, "Monotone control systems," *IEEE Transactions on automatic control*, vol. 48, no. 10, pp. 1684–1698, 2003.
- [33] H. Klärner, A. Bockmayr, and H. Siebert, "Computing symbolic steady states of boolean networks," in *Proceedings of the 11th International Conference on Cellular Automata for Research and Industry, ACRI 2014*, ser. Lecture Notes in Computer Science, J. Waş, G. C. Sirakoulis, and S. Bandini, Eds., vol. 8751. Springer-Verlag, 2014, pp. 561–570.
- [34] S. Gay, S. Soliman, and F. Fages, "A graphical method for reducing and relating models in systems biology," *Bioinformatics*, vol. 26, no. 18, pp. i575–i581, 2010, special issue ECCB'10.
- [35] J.-C. Leloup and A. Goldbeter, "Toward a detailed computational model for the mammalian circadian clock," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 7051–7056, 2003.
- [36] H. Klärner, A. Bockmayr, and H. Siebert, "Computing maximal and minimal trap spaces of boolean networks," *Natural Computing*, vol. 14, no. 4, pp. 535–544, 2015.



**François Fages** François Fages is Research Director (DR1) at Inria where he leads the Lifeware group and coordinates the development of the BIOCHAM modeling software for systems biology. In 1983, he passed his Doctorate Thesis on unification theory, and got a research position at CNRS, Ecole Normale Supérieure, with part-time assistant professorship at Ecole Polytechnique (1985–1998) and part-time consultancy at Thomson-CSF/Thales group (1985–1996). In 1999 he joined Inria and started working on

computational systems biology where he has now published 70 articles over a total of more than 200 papers and 3700 citations. He has supervised 29 PhD Theses. He collaborates with biologists on challenging questions in cell signaling and cell cycle regulation, and leads a Master course on “Computational methods for systems and synthetic biology”. In 2014, he received the Michel Monpetit prize from the French Academy of Sciences.



**Thierry Martinez** Thierry Martinez is research engineer (IR2) at Inria de Paris. He passed his Ph.D. in 2015 on constraint and logic programming, in the Lifeware team. He has been contributing to the BIOCHAM modelling platform, during his Ph.D. and since then, as research associate. At Inria de Paris, he has also been contributing to Coccinelle (a software to check and strengthen the quality of the Linux kernel) and Coq (a formal proof assistant). He has published about 30 papers (cf. Google Scholar), on

constraint and logic programming, but also complexity theory, graph theory, etc.



**David Rosenblueth** David A. Rosenblueth graduated at the University of Victoria, B. C., Canada. He has worked in logic programming, model checking and complex systems. Within logic programming, he has contributions to inference systems inspired by parsers, inductive logic programming, program transformation, and logic programming applied to genetic regulation. Within model checking, he has studied the model-update problem and model checking for hidden Markov models, multi-agent systems, genetic regulation, robotics, and embedded systems. Within complex systems, he has worked in self-organizing traffic lights.

genetic regulation, robotics, and embedded systems. Within complex systems, he has worked in self-organizing traffic lights.



**Sylvain Soliman** Sylvain Soliman is permanent researcher (CR1) at Inria Saclay-Ile-de-France, in the Lifeware team. He passed his Ph.D. in 2001 on Linear Logic, did a post-doc in a DGA lab on Natural Language Processing and then was hired at Inria in 2002. Since then he has been working and teaching Constraint Programming, and from 2004 has been one of the core developers of the BIOCHAM modelling platform, therefore moving to the field of Computational Systems Biology. Using methods from theoretical computer science, like constraint programming but also model checking, abstract interpretation, etc. he has published about 100 papers (cf. Google Scholar) and co-supervised 2 Ph.D. theses. He obtained his Habilitation degree (HDR) in 2016.

ical computer science, like constraint programming but also model checking, abstract interpretation, etc. he has published about 100 papers (cf. Google Scholar) and co-supervised 2 Ph.D. theses. He obtained his Habilitation degree (HDR) in 2016.