

# Systemic approach for toxicity analysis

Cinzia Di Giusto, Hanna Klaudel, and Franck Delaplace

Université d'Evry - Val d'Essonne, Laboratoire IBISC, Evry, France

**Abstract.** A high-level Petri net framework is introduced for the toxic risk assessment in biological and bio-synthetic systems. Unlike empirical techniques mostly used in toxicology or toxicogenomics, we propose a systemic approach consisting of a series of behavioral rules (reactions) that depend on abstract discrete “expression” levels of involved agents (species). We introduce a finite state high-level Petri net model allowing exhaustive verification (model-checking) of properties related to equilibrium alteration or appearing of hazardous behaviors. The approach is applied to the study of the impact of the aspartame assimilation into the blood glucose regulation process.

## 1 Introduction

Toxicology [23] studies the adverse effects of the exposures to chemicals at various levels of living entities: organism, tissue, cell or intracellular molecular systems. During the last decade, the accumulation of genomic and post-genomic data together with the introduction of new technologies for gene analysis has opened the way to *toxicogenomics*. Toxicogenomics combines toxicology with “Omics” technologies<sup>1</sup> to study the *mode-of-action* of toxicants or environmental stressors on biological systems. The mode-of-action is understood as the sequence of events from the absorption of chemicals to a toxic outcome. Toxicogenomics potentially improves clinical diagnosis capabilities and facilitates the identification of potential toxicity in drug discovery [10] or in the design of bio-synthetic entities [21].

The main approach used in toxicogenomics employs empirical analysis like in the identification of molecular biomarkers, i.e., indicators of disease or toxicity in the form of specific gene expression patterns [7]. Clearly, biomarkers remain observational indicators linking genes related measures to toxic states. In this proposal, we complement these empirical methods with a computational technique that aims at discovering the molecular mechanisms of toxicity. This way, instead of studying the phenomenology of the toxic impacts, we focus on the processes triggering adverse effects on organisms. Usually, the toxicity process is defined as a sequence of physiological events that causes the abnormal behavior of a living organism with respect to its healthy state. Healthy physiological states generally correspond to homeostasis, namely a process that maintains a dynamic stability of internal conditions against changes in the external environment. Hence, we

---

<sup>1</sup> “Omics” technologies are methodologies such as genomics, transcriptomics, proteomics and metabolomics.

will consider toxicity outcomes as deregulation of homeostasis processes, namely deviation of some intrinsic referential equilibrium of the system.

Biological processes are usually given in terms of pathways which are causal chains of the responses to stimuli, this way the deregulation of homeostasis appears as the activation or inhibition of unexpected but existing pathways. Moreover, in the context of toxicogenomics it is crucial to take into account at least two other parameters: the exposure time and the thresholds dosage delimiting the ranges of safe and hazardous effects.

In this paper, we depict and analyze the mechanistic process of toxicology using high-level Petri nets. Our work is inspired by the definition of reaction systems as given in [1]. A reaction system is a set of *reactions*, each of them defined as a triple  $(R, I, P)$  where  $R$  is the set of reactants,  $I$  the set of inhibitors and  $P$  the set of products, and  $R, I$  and  $P$  are taken from a common set of species  $S$ . Reaction systems are based on three foundational principles:

1. a reaction can take place only if all the reactants involved are available but none of the inhibitors is;
2. if a species is available then a sufficient amount of it is present to trigger a reaction;
3. species are not persistent: they become unavailable if they are not sustained by a reaction.

From this model we retain the idea of reactions but we significantly change the semantics. The first change concerns principle 2: species are available at a given discrete abstract level. This is mainly related to the need of expressing toxicants doses. The corresponding discretization is built observing thresholds levels in dose-response curves. The second and more fundamental change regards the introduction of discrete time constraints. Time plays a role in the evolution of species, more precisely, species are associated to a decay time  $\delta$ , meaning that their level diminishes with time. This accounts for the presence of a non-specified environment that consumes and degrades species, thus allowing to abstract away from reactions that may be neglected in the specified context. Each reaction  $(R, I, P)$  is extended with levels for all its reactants and inhibitors. Reactions can take place only if each reactant is present at least at a given level and each involved inhibitor is at a level strictly inferior to the given one. As a result, the level of products of the reaction can be increased or decreased.

Summing up, systems are build out of a series of behavioral reactions among involved agents or species. We model such systems into high-level Petri nets and apply it to toxicogenomics problems, namely deregulation of homeostatic processes. Toxicity questions are expressed using a suitable temporal logic like CTL [9]. By observing that our modeling has a finite state space, it is therefore natural to address the satisfiability of these formulae using classic verification techniques such as model checking.

We apply the modeling and verification process on the example of blood glucose regulation in human body showing the maintenance of the homeostasis. In particular, we highlight how the interplay between the assimilation of aspartame and glucose regulation causes the appearance of unwanted behaviors.

**Organization of the paper.** The paper is organized as follows: Section 2 recalls basic definitions and notations on high-level Petri nets. Next, Section 3 describes our running example of blood glucose regulation. Section 4 introduces the principles behind reaction networks and presents their high-level Petri net modeling. Then Section 5 shows how to check toxicology properties and finally, Section 7 concludes with some considerations on future work.

## 2 Preliminaries

We recall here the general notations together with some elements of the semantics of high-level Petri nets [15].

**Definition 1.** A high-level Petri net  $N$  is a tuple  $(Q, T, F, L, M_0)$  where:

- $Q$  is the set of places,
- $T$  is the set of transitions and  $Q \cap T = \emptyset$ ;
- $F \subseteq (Q \times T) \cup (T \times Q)$  is the set of arcs;
- $L$  is the labeling function from places  $Q$ , transitions  $T$  and arcs  $F$  to a set of labels defined as follows:
  - $\forall q \in Q$ ,  $L(q)$  is the type of  $q$ , i.e., a (possibly infinite) set or Cartesian product of sets of integer values;
  - $\forall t \in T$ ,  $L(t)$  is a computable boolean expression with variables and integer values;
  - and  $\forall f \in F$ ,  $L(f)$  is a tuple of variables and integer values compatible with the adjacent place.
- $M_0$  is the initial marking which associates to each place  $q \in Q$  a multiset of tokens in  $L(q)$ .

Observe that we are considering a subclass of high-level Petri nets where at most one arc per direction for each pair place/transition is allowed and only one token can flow through. The behavior of high-level Petri nets is defined as usual: markings are functions from places in  $Q$  to multisets of possibly structured tokens in  $L(q)$  and a transition  $t \in T$  is *enabled* at marking  $M$ , if there exists an evaluation  $\sigma$  of all variables in the labeling of  $t$  such that the guard  $L(t)$  evaluates to true ( $L_\sigma(t) = \text{true}$ ) and there are enough tokens in all input places  $q$  to satisfy the corresponding input arcs, i.e.,  $L_\sigma((q, t)) \in M(q)$ . Then, the firing of  $t$  produces the marking  $M'$ :

$$\forall q \in Q, M'(q) = M(q) - L_\sigma((q, t)) + L_\sigma((t, q)).$$

with  $L_\sigma(f) = 0$  if  $f \notin F$ ,  $-$  and  $+$  are multiset operators for removal and adding of one element, respectively. We denote it by  $M[t:\sigma]M'$ .

By convention, primed version of variables (e.g.  $x'$ ) are used to annotate output arcs of transitions, their evaluation is possibly computed using unprimed variables (e.g.  $x$  and  $y$ ) appearing on input arcs. With an abuse of notation, singleton markings are denoted without brackets, the same is used in arc annotations. An example of firing is shown in Figure 1. We say that a marking  $M$  is *reachable* from the initial marking  $M_0$  if there exists a firing sequence  $(t_1, \sigma_1), \dots, (t_n, \sigma_n)$  such that  $M_0[t_1:\sigma_1]M_1 \dots M_{n-1}[t_n:\sigma_n]M$ .

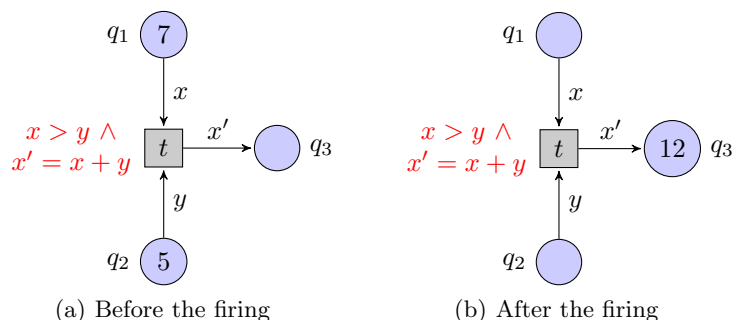


Fig. 1. Example of firing with  $\sigma = \{x = 7, y = 5, x' = 12\}$ .

### 3 Blood glucose regulation

Here we introduce our running example: *glucose regulation* in human body (Figure 2). In the following, we are always referring to the process under normal circumstances in a healthy body.

Glucose regulation is a homeostatic process: i.e., the rates of glucose in blood (*glycemia*) must remain stable at what we call the equilibrium state. Glycemia is regulated by two hormones: *insulin* and *glucagon*. When glycemia rises (for instance as a result of the digestion of a meal), insulin promotes the storing of glucose in muscles through the glycogenesis process, thus decreasing the blood glucose levels. Conversely, when glycemia is critically low, glucagon stimulates the process of glycogenolysis that increases the blood glucose level by transforming glycogen back into glucose.

We will focus on the assimilation of sweeteners: i.e., sugars or artificial sweeteners such as aspartame. Whenever we eat something sweet either natural or artificial, the sweet sensation sends a signal to the brain (through *neurotransmitters*) that in turns stimulates the production of insulin by pancreas. In the case of sugar, the digestion transforms food into nutrients (i.e., glucose) that

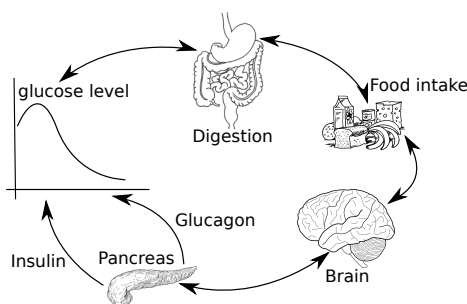


Fig. 2. Glucose metabolism

are absorbed by blood. This way, sugar through digestion increases glucose in blood giving the sensation of satiety. In case the income of glucose produces hyperglycemia, the levels of glucose are promptly equilibrated by the intervention of insulin. Unlike sugar, artificial sweeteners are not assimilated by the body, hence they do not increase the glucose levels in blood. Nevertheless the insulin produced under the stimuli originated by the sweet sensation, although weak, can still cause the rate of glucose to drop engendering hypoglycemia. In response to that, the brain induces the stimulus of *hunger*. As a matter of fact this appears as an unwanted/toxic behavior, indeed the assimilation of food (even if it contains aspartame) should calm hunger and induce satiety not the opposite.

This schema suggests that we should consider four levels for glycemia: low, hunger, equilibrium and high. Likewise for insulin we assume three levels: inactive, low and high. All other actors involved in glucose regulation, have only two levels (inactive or active). In the following sections, we will see how to model the glucose metabolism and how to verify the unexpected behaviors of artificial sweeteners.

## 4 Petri net modeling

A reaction network is composed of a set of species  $\mathcal{S}$  governed by a set of reactions  $\mathcal{R}$ . We begin by giving some intuitions on their dynamics.

Species in  $\mathcal{S}$  represent the actors of the modeled system. In the example introduced above, we have concrete species such as aspartame and also more abstract ones representing ratios or concepts like glycemia. Species may have several *expression* levels. Levels are determined by the observable behavior of species, i.e., they refer to a change in the capability of action of species. In toxicology, they may represent dosages. We assume, for each species  $s$ , an arbitrary but finite number  $\mathcal{L}_s$  of levels, and each  $s$  is initialized at a certain level  $\eta_s$ . For certain species, we assume the presence of a non specified environment that acts on them by decreasing gradually their expression levels. This special activity is called *decay* and is modeled by various durations associated to expression levels. Decay may be unbounded indicating that the level of the species can only change by result of a reaction. It is formalized by a function that associates to each level either  $\omega$  (unbounded) or its finite duration:

$$\delta_s : [0.. \mathcal{L}_s - 1] \rightarrow \mathbb{N}^+ \cup \{\omega\}.$$

For all species  $s \in \mathcal{S}$  we require that  $\delta_s(0) = \omega$  meaning that the duration of the basal level must be unbounded.

*Example 1 (Glucose metabolism – species).* Take the example from Section 3. The set of involved species is

$$\mathcal{S} = \{Sugar, Aspartame, Glycemia, Glucagon, Insulin\}$$

and their expression levels and corresponding decays are:

levels	durations
$\mathcal{L}_{sugar} = \{0, 1\}$	$\delta_{sugar}(1) = 2$
$\mathcal{L}_{aspartame} = \{0, 1\}$	$\delta_{aspartame}(1) = 2$
$\mathcal{L}_{glycemia} = \{0, 1, 2, 3\}$	$\delta_{glycemia}(1) = 8$ $\delta_{glycemia}(2) = 8$ $\delta_{glycemia}(3) = 8$
$\mathcal{L}_{glucagon} = \{0, 1\}$	$\delta_{glucagon}(1) = 3$
$\mathcal{L}_{insulin} = \{0, 1, 2\}$	$\delta_{insulin}(1) = 3$ $\delta_{insulin}(2) = 3$

The levels of glycemia are: 0 corresponding to low, 1 to hunger, 2 to equilibrium and 3 to high. Likewise for insulin we have 0 that corresponds to inactive, 1 to low and 2 to high. All levels for the other species are 0 for inactive and 1 for active.  $\diamond$

The evolution of species  $s \in \mathcal{S}$  is governed by a set of reactions  $\mathcal{R}$ , each being of the form:

$$\rho ::= \langle R_\rho, I_\rho, P_\rho \rangle \quad (1)$$

where  $R_\rho$  (reactants),  $I_\rho$  (inhibitors) are sets of pairs  $(s, \eta_s)$  and  $P_\rho$  (products) is a non empty set of pairs  $(s, z)$ , where  $\eta_s \in [0..|\mathcal{L}_s - 1]$  and  $z \in \mathbb{Z}$ . Species can appear at most once in each set  $R_\rho$ ,  $I_\rho$  and  $P_\rho$ . They can be present in both  $R_\rho$  and  $I_\rho$  but they must occur with different levels<sup>2</sup>. We write  $s \in R_\rho$  to denote  $(s, \cdot) \in R_\rho$  similarly for  $I_\rho$  and  $P_\rho$  and we omit index  $\rho$  if it is clear from the context ( $\rho = \langle R, I, P \rangle$ ).

*Example 2 (Glucose metabolism – reactions).* The set of reactions  $\mathcal{R} = \{\rho_k = (R_k, I_k, P_k) \mid k \in [1..9]\}$  for the glucose metabolism example is:

$\rho_k$	Reactants $R_k$	Inhibitors $I_k$	Products $P_k$
$\rho_1$	(Sugar, 1)	$\emptyset$	(Insulin, +1), (Glycemia, +1)
$\rho_2$	(Aspartame, 1)	$\emptyset$	(Insulin, +1)
$\rho_3$	$\emptyset$	(Glycemia, 1)	(Glucagon, +1)
$\rho_4$	(Glycemia, 3)	$\emptyset$	(Insulin, +1)
$\rho_5$	(Insulin, 2)	$\emptyset$	(Glycemia, -1)
$\rho_6$	(Insulin, 1), (Glycemia, 3)	$\emptyset$	(Glycemia, -1)
$\rho_7$	(Insulin, 1)	(Glycemia, 2)	(Glycemia, -1)
$\rho_8$	(Glucagon, 1)	$\emptyset$	(Glycemia, +1)

$\rho_1$  and  $\rho_2$  represent the assimilation of Sugar and Aspartame, respectively: while Aspartame only increases the level of Insulin, Sugar also increases Glycemia.  $\rho_3$  takes care of hypoglycemia, i.e., a Glycemia level equal to 0 (obtained by

<sup>2</sup> Observe that a species can appear in the same reaction as reactant at level  $\eta_r$ , inhibitor at level  $\eta_i > \eta_r$  and product.

using (*Glycemia*, 1) as inhibitor) engenders the production of Glucagon. On the contrary, hyperglycemia causes the production of Insulin ( $\rho_4$ ). The presence of Insulin lowers Glycemia (reactions  $\rho_5, \rho_6, \rho_7$ ). In particular Insulin level equal to 1 plays a role in the decrease of Glycemia only in case of hyperglycemia  $\rho_6$  or hypoglycemia  $\rho_7$ , otherwise the signal is not strong enough and we need Insulin at level 2 to see the effect on Glycemia ( $\rho_5$ ). Last reaction describes the role of Glucagon which if active increases the level of Glycemia.

◇

The dynamics of reaction networks is formalized using high-level Petri nets. We represent the state of a species  $s$  as a pair  $\langle l_s, u_s \rangle$ , where  $l_s$  is an integer value storing the current level from zero to  $\mathcal{L}_s - 1$ , and  $u_s$  is a counter storing the interval of time spent at level  $l_s$ . The system is initialized by setting the level of all species: i.e., each species  $s$  is set to  $\langle \eta_s, 0 \rangle$  where  $\eta_s$  is the given initial level.

Reaction networks can evolve in two ways:

- Case 1. **Time progression and Decay:** Time progresses discretely of one unit at once. It affects species with finite decay only. More precisely if a species  $s$  has unbounded decay at level  $l$  ( $\delta_s(l) = w$ ) then its corresponding tuple  $(\eta_s, u_s)$  remains unchanged. Otherwise, if the species has a finite decay ( $\delta_s(l) = d$ ), it may stay at level  $l$  for  $d$  time units. Then, degradation happens as soon as  $d$  time units are elapsed and is obtained by decreasing the level to  $l - 1$  and by setting  $u_s$  to zero.
- Case 2. **Reaction:** A reaction  $\rho$  may happen if and only if all the reactants are available at least at the required level and all the inhibitors are expressed at a level strictly inferior to the required one. The triggering of a reaction results in the update of the level of all its products. Depending on the reaction, levels will be increased (+n), maintained (0) or decreased (-n). We assume that each reaction can take place only once per time unit.

We now comment on some specific design choices concerning reactions:

- the set of reactants and inhibitors  $R \cup I$  is allowed to be empty. This accounts for modeling an environment that is continuously sustaining the production of a species.
- a species can appear in the same reaction simultaneously as a reactant and an inhibitor. In such a case, we require them to occur with different levels:

$$(\{(s, \eta)\} \cup R, \{(s, \eta')\} \cup I, P)$$

where  $\eta < \eta'$ . This means that the reaction can take place only if the level  $l_s$  of  $s$  belongs to the interval  $\eta \leq l_s < \eta'$ . In particular, if  $s$  has to be present in a reaction exactly at level  $\eta$ ,  $s$  should appear as a reactant at level  $\eta$  and as inhibitor at level  $\eta' = \eta + 1$ ;

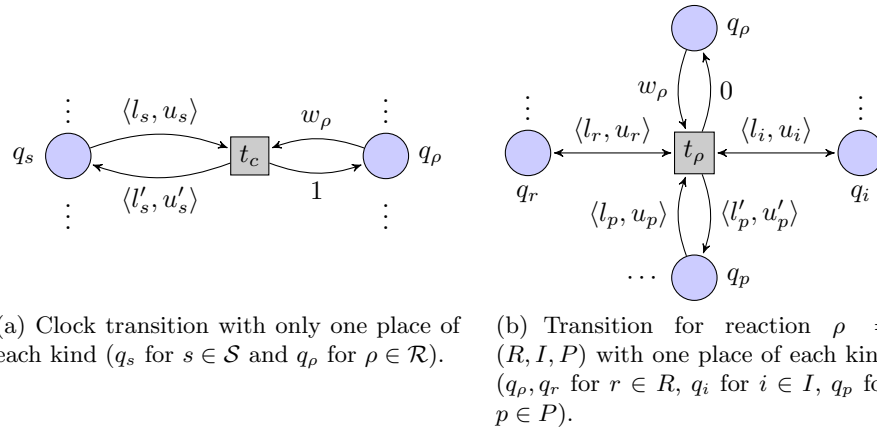
- species can appear only once in the set of products  $P$ . This implies that a product cannot be increased and decreased in the same reaction.

It is also worth observing that if a species is continuously sustained by some reactions then it remains available in the system at a certain level for a period that could be longer than the corresponding decay time.

*Example 3 (Glucose metabolism – scenario).* Take once again the example of glucose metabolism and observe the behavior of *Glycemia* in the following scenario:

initial state	$\langle 3, 0 \rangle$
8 time units elapse, counter at level 3 updates	$\langle 3, 8 \rangle$
one time unit elapses, <i>Glycemia</i> decays	$\langle 2, 0 \rangle$
one time unit elapses, counter at level 2 updates	$\langle 2, 1 \rangle$
reaction $\rho_5$ decreases <i>Glycemia</i> level	$\langle 1, 0 \rangle$
8 time units elapse, counter at level 1 updates	$\langle 1, 8 \rangle$
one time unit elapses, <i>Glycemia</i> decays	$\langle 0, 0 \rangle$
one time unit elapses, no effect since $\delta_{glycemia}(0) = \omega$	$\langle 0, 0 \rangle$ . $\diamond$

More formally, we now introduce the high-level Petri net modeling. Each species  $s \in \mathcal{S}$  is modeled by a single place  $q_s$  whose type  $L(q_s)$  is the set of tuples of the form  $\langle l_s, u_s \rangle$ , where  $l_s \in [0..L_s - 1]$  and  $u_s \in [0..max_s]$ , with  $max_s = \max\{\delta_s(l) \mid \delta_s(l) \neq \omega \text{ and } l \in [0..L_s - 1]\}$ . In order to cope with time aspects we introduce a transition  $t_c$  (Figure 3(a)) connected to all species that is responsible for time progression and takes care of the decay of concerned species (as described in Case 1 above). Finally, every reaction  $\rho$  is modeled with a transition  $t_\rho$  (Figure 3(b)). To each transition  $t_\rho$  we associate a special place  $q_\rho$  that is used to ensure that the same reaction is not executed more than once in the same time unit. More detailed explanations for each type of transition follow Definition 2.



**Fig. 3.** Scheme of Petri net modeling of reaction networks.

**Definition 2.** Given a network  $(\mathcal{S}, \mathcal{R})$  with initial state  $(s, \eta_s)$  for each  $s \in \mathcal{S}$ , its high-level Petri net representation is defined as tuple  $(Q, T, F, L, M_0)$  where  $z, z', l, l', u, u', w, w'$  are variables and:



- $Q = \{q_s \mid s \in \mathcal{S}\} \cup \{q_\rho \mid \rho \in \mathcal{R}\}$ ;
- $T = \{t_c\} \cup \{t_\rho \mid \rho \in \mathcal{R}\}$ ;
- $F = \{(q, t_c), (t_c, q) \mid q \in Q\} \cup$   
 $\{(q_s, t_\rho), (t_\rho, q_s), (q_\rho, t_\rho), (t_\rho, q_\rho) \mid \rho \in \mathcal{R}, s \in R_\rho \cup I_\rho \cup P_\rho\}$
- Labels for places in  $Q$ :

$$L(q_\rho) = \{0, 1\} \text{ for each } \rho \in \mathcal{R}$$

$$L(q_s) = [0..L_s - 1] \times [0..max_s] \text{ for each } s \in \mathcal{S}$$

- Labels for arcs in  $F$ :

$$L((q_\rho, t_c)) = w \quad L((t_c, q_c)) = 1$$

$$L((q_s, t_c)) = \langle l_s, u_s \rangle \quad L((t_c, q_s)) = \langle l'_s, u'_s \rangle \text{ for each } s \in \mathcal{S}$$

For each reaction  $\rho \in \mathcal{R}$  and  $s \in R_\rho \cup I_\rho \cup P_\rho$ :

$$L((q_s, t_\rho)) = \langle l_s, u_s \rangle \quad L((t_\rho, q_s)) = \begin{cases} \langle l_s, u_s \rangle & \text{if } s \notin P_\rho \\ \langle l'_s, u'_s \rangle & \text{otherwise} \end{cases}$$

$$L((q_\rho, t_\rho)) = w \quad L((t_\rho, q_\rho)) = 0$$

- Labels for transitions in  $T$ :

$$L(t_c) = \bigwedge_{s \in \mathcal{S}} ((\delta(l_s) = \omega \vee u_s + 1 \leq \delta(l_s)) \rightarrow \langle l'_s, u'_s \rangle = \langle l_s, u_s + 1 \rangle \wedge$$

$$(\delta(l_s) \neq \omega \wedge u_s + 1 > \delta(l_s)) \rightarrow \langle l'_s, u'_s \rangle = \langle l_s - 1, 0 \rangle).$$

For each reaction  $\rho \in \mathcal{R}$ :

$$L(t_\rho) = (w = 1) \wedge \bigwedge_{(r, \eta_r) \in R_\rho} (l_r \geq \eta_r) \wedge \bigwedge_{(i, \eta_i) \in I_\rho} (l_i < \eta_i) \wedge$$

$$\bigwedge_{(p, z) \in P_\rho} (\langle l'_p, u'_p \rangle = \langle \max(0, \min(l_p + z, L_p - 1)), 0 \rangle)$$

- For each  $q \in Q$ ,  $s \in \mathcal{S}$  and  $\rho \in \mathcal{R}$ , the initial marking  $M_0$  is:

$$M_0(q) = \begin{cases} 1 & \text{if } q = q_\rho, \\ \langle \eta_s, 0 \rangle & \text{if } q = q_s. \end{cases}$$

We now comment on the transitions of the high-level Petri net. The result of the firing of a transition is handled by guards (namely transition labels  $L(t_c)$  and  $L(t_\rho)$ ) together with the evaluation  $\sigma$  as described after Definition 1. With an abuse of notation, in the following, we refer to evaluated variables without effectively mentioning the evaluation  $\sigma$ : i.e., we say that the current value of the token in  $q_\rho$  is  $w$  instead of  $\sigma(w)$ . Input and output arcs between the same place and transition with the same label (read arcs) are denoted in figures with a double-pointed arrow with a single label.

Clock transition  $t_c$ , depicted in Figure 3(a), takes care of Case 1 above.  $t_c$  is responsible for the decay of concerned species and the related update of counters  $u_s$  of each species. Moreover  $t_c$  updates the tokens of all places  $q_\rho$  to 1, thus re-enabling the possibility of performing a reaction  $\rho$ .

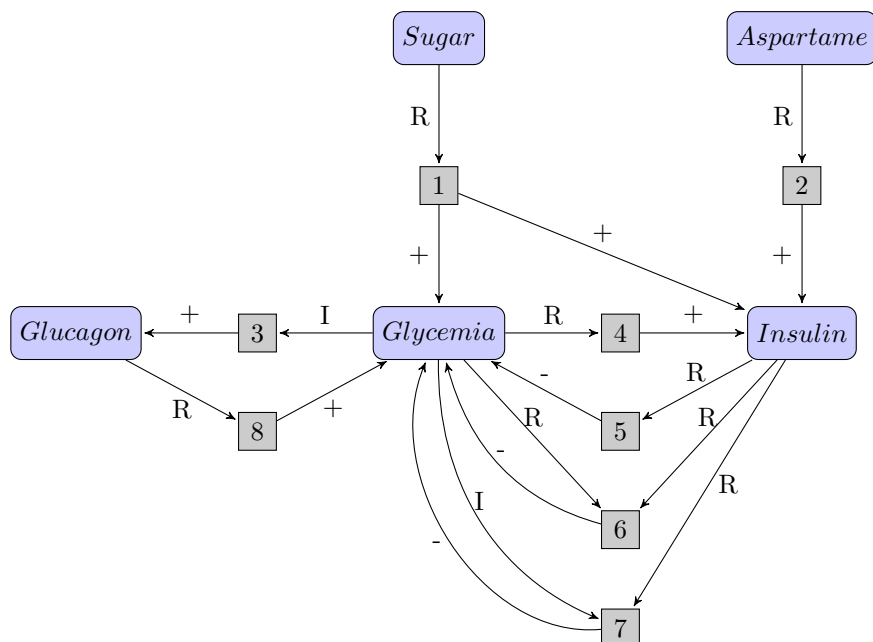
Next, we describe transitions for reactions, depicted in Figure 3(b). Given a reaction  $\rho = (R, I, P)$  we detail the conditions and the results of firing of  $t_\rho$ . As described in Case 2 we have:

- each reactant  $r \in R$  has to be present at least at level  $\eta_r$ , this is expressed by guard  $l_r \geq \eta_r$ ;
- each inhibitor  $i \in I$  has not to exceed level  $\eta_i$ , this is guaranteed by guard ( $l_i < \eta_i$ );
- each product  $p \in P$  corresponding to place  $q_p$  is updated to  $\langle l'_p, u'_p \rangle = \langle \max(0, \min(l_p + z, \mathcal{L}_p - 1)), 0 \rangle$ .

The role of place  $q_p$  is to forbid two consecutive executions of the same reaction in the same time unit. Initially, the marking of  $q_p$  is set to 1 and it becomes 0 when the transition  $t_p$  is fired; then clock transition  $t_c$  sets it to 1 again.

Observe that, because of the semantics of high level Petri nets, reaction may not occur even if all constraints are satisfied. This is interpreted as the action of an hostile (non-specified) environment (e.g., reactants are too far from each other to react).

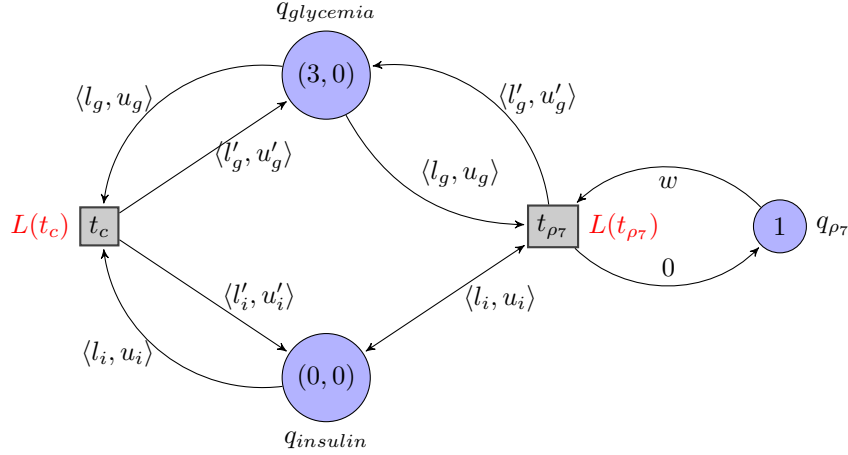
*Example 4 (Glucose metabolism – reaction network).*



**Fig. 4.** Simplified reaction network of glucose metabolism.

Figure 4 shows a simplification of the reaction network  $(\mathcal{S}, \mathcal{R})$  given in example 1. It focuses only on the reaction schema linking inputs (i.e., reactants and inhibitors) to products. Each input arc is labeled with either letter R or letter I denoting whether the input place is a reactant or an inhibitor, respectively. Likewise, each output arc is labeled with a + or a - to denote increase or decrease

of product levels by 1. For each reaction transition  $\rho$ , we have omitted place  $q_\rho$  and all arcs in the opposite direction. The numbers inside each transition refers to the corresponding reaction in Example 2.



**Fig. 5.** A portion of the reaction network of glucose metabolism with an initial marking.

Figure 5, instead, shows a portion of the complete initially marked reaction network for the glucose metabolism example, focusing only on reaction  $\rho_7$ .  $\diamond$

From the above definition and the transition rule of high-level nets, we have the following properties:

**Proposition 1.** *Given a reaction network  $(\mathcal{S}, \mathcal{R})$  with initial values  $\eta_s$  for each species  $s \in \mathcal{S}$ :*

- its Petri net representation has a finite structure with  $|\mathcal{S}| + |\mathcal{R}|$  places,  $|\mathcal{R}| + 1$  transitions and the number of arcs is bounded by  $2(|\mathcal{S}| + |\mathcal{R}| + \sum_{\rho=(R_\rho, I_\rho, P_\rho) \in \mathcal{R}} (|R_\rho| + |I_\rho| + |P_\rho| + 1))$ ;
- each place type is a finite set;
- for each arc  $(q, t) \in F$  there is an arc in opposite direction, i.e.,  $(t, q) \in F$  and each arc label is a singleton;
- the initial marking and all reachable markings have exactly one structured token per place;
- the number of all reachable markings from the initial one is finite.

*Proof.* Follows by definition and by induction on the length of a firing sequence.  $\square$

## 5 Toxicology analysis

Such a Petri net representation of a reaction network is used to detect and predict toxic behaviors related to the dynamics of bio-molecular networks. In order to verify toxicology properties, we resort to temporal logics and model checking techniques [5]. For the sake of the present paper computation tree logic (CTL) allows to express properties of interest. Nonetheless different scenarios may require other more appropriate modal logic which we could be handled by our framework.

We recall here the basic concepts of CTL, provide the formal definition of the syntax and give some intuitions on the semantics, formally defined in [9].

A CTL formula is defined as:

$$\varphi ::= \perp \mid a \mid \neg\varphi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \varphi \rightarrow \varphi \\ \mathbf{EX}\varphi \mid \mathbf{EG}\varphi \mid \mathbf{E}(\varphi\mathbf{U}\varphi) \mid \mathbf{EF}\varphi \mid \mathbf{AG}\varphi \mid \mathbf{AF}\varphi$$

where  $a \in A$  is an atomic proposition.

CTL is used to state properties on branching time structures. It uses usual boolean operators, path quantifiers and temporal operators. Path quantifiers can be of two kinds:  $\mathbf{A}\varphi$  means that  $\varphi$  has to hold on all paths starting from the current state, while  $\mathbf{E}\varphi$  stands for there exists at least one path starting from the current state where  $\varphi$  holds. We have four temporal operators:  $\mathbf{X}\varphi$  holds if  $\varphi$  is true at the *next* state,  $\mathbf{G}\varphi$  means that  $\varphi$  has to *globally* hold on the entire subsequent path,  $\mathbf{F}\varphi$  stands for eventually (or *finally*)  $\varphi$  has to hold (at some point on the subsequent path), and  $\varphi_1\mathbf{U}\varphi_2$  means that  $\varphi_1$  has to hold at least *until* at some position  $\varphi_2$  holds. In our context atomic formulae are represented by pairs of species and levels:  $A = \{(s, \eta_s) \mid s \in \mathcal{S}\}$ , for instance  $(Glucose, 2)$ .

As mentioned in the introduction, we are mainly interested in checking whether the inner equilibrium of an organism (tissue, cell, ...) is maintained when administrating drugs or applying stressors. More in detail, toxicology properties can be classified into two categories:

- properties checking for the appearance of particular symptoms, and
- properties characterizing causal relations between events.

The former class of properties basically consists in verifying reachability of some states, while the latter concerns pathways that highlight sequences of events leading to toxic outcomes. For instance, in the case of glucose regulation, we could verify whether glycemia levels are kept stable and whether they change in case of ingestion of aspartame. More precisely, we could examine the causes and the symptoms of the hypoglycemia induced by the assimilation of aspartame. Hence hypoglycemia is treated as a toxic state.

*Example 5 (Glucose metabolism – properties).* Take our running example of blood glucose regulation. The following properties can be expressed in CTL:

**Symptoms:** Is it possible to have an anomalous decrease of glucose levels in blood (revealing hypoglycemia)?

$$\mathbf{EF}(Glycemia, 0)$$

**Mode-of-action:** Recalling that the blood glucose regulation process normally maintains glycemia at equilibrium (level 2), is there an abnormal behavior leading to hypoglycemia?

$$\mathbf{E}(\mathbf{EF}(\textit{Glycemia}, 2) \mathbf{U} (\mathbf{EF}(\textit{Glycemia}, 0)))$$

**Causality:** Does assimilation of sweeteners cause hypoglycemia?

$$\mathbf{EF}[\textit{Sugar}, 1) \vee (\textit{Aspartame}, 1)) \wedge (\textit{Glycemia}, 1)] \rightarrow \mathbf{AF}(\textit{Glycemia}, 2)$$

For the third formula we show two paths given as sequences of reactions (abstracting away from time transitions), one that satisfy the formula and the other that contradicts it. The first one corresponds to the assimilation of sugar. As described in Section 3, the digestion of sugar induces an increase of the production of insulin and an augmentation of the blood glucose levels. Nonetheless the levels of insulin produced are not enough to cause the glycemia to drop and the formula is satisfied.

$$\begin{aligned} & (\textit{Sugar}, 1), (\textit{Aspartame}, 0), (\textit{Glycemia}, 1), (\textit{Insulin}, 0), (\textit{Glucagon}, 0) \xrightarrow{\rho^1} \\ & (\textit{Sugar}, 1), (\textit{Aspartame}, 0), (\mathbf{Glycemia}, 2), (\textit{Insulin}, 1), (\textit{Glucagon}, 0) \end{aligned}$$

Unlike previous path, the assimilation of aspartame causes only an increase of insulin. Unfortunately, this increment is sufficient to induce a decrease of blood glucose levels thus contradicting the formula above.

$$\begin{aligned} & (\textit{Sugar}, 0), (\textit{Aspartame}, 1), (\textit{Glycemia}, 1), (\textit{Insulin}, 0), (\textit{Glucagon}, 0) \xrightarrow{\rho^2} \\ & (\textit{Sugar}, 0), (\textit{Aspartame}, 1), (\textit{Glycemia}, 1), (\textit{Insulin}, 1), (\textit{Glucagon}, 0) \xrightarrow{\rho^3} \\ & (\textit{Sugar}, 0), (\textit{Aspartame}, 0), (\mathbf{Glycemia}, 0), (\textit{Insulin}, 1), (\textit{Glucagon}, 0) \end{aligned}$$

This illustrates the toxic behavior caused by aspartame described in Section 3.  $\diamond$

## 6 Related work

The main application of our work concerns the verification of properties of systems defined in terms of rules or reactions. From a technical point of view, the closest related work is on reaction systems [1] or their Petri net representation [17]. Although we use a similar definition for reactions, the semantics that we have proposed is inherently different: in [1] all enabled reactions occur in one step while we have considered an interleaving semantics. In [2], the authors consider an extension of reaction systems with a notion of decay, this concept is different from the one considered here as we refer to an independent time progression while they count the number of maximally concurrent steps. In fact, our representation of time is considerably different from the approaches traditionally used in time and timed Petri nets ([3] presents a survey with insightful comparison of the different approaches). The main difference lies on the fact that the progression of time is implicit and external to the system. By contrast, in

our proposal we have assumed the presence of an explicit way of incrementing durations (modeled by synchronized counters). This is also different from the notion of timestamps introduced in [12] that again refers to an implicit notion of time. Indeed, our approach is conceptually closer to Petri nets with causal time [22] for the presence of an explicit transition for time progression. Nevertheless, in reaction networks time cannot be suspended under the influence of the environment (as is the case in [22]).

In a broader sense, our work could also be related to P-systems [18,16] or the  $\kappa$ -calculus [6] that describe the evolution of cells through rules. Both these approaches are mainly oriented to simulation while we are interested in verification aspects. Finally, always related to the modeling in Petri nets but with a different aim, levels have been used in qualitative approaches to address problems related to the identification of steady states in genetic networks such as in [4]. Nevertheless these contributions abstract away from time related aspects that are instead central in our proposal.

## 7 Conclusion and future work

We have introduced a high-level Petri net modeling of reaction networks to address problems related to toxicogenomics. In reaction networks, systems consist of a set of species present in the environment at a given level. Species can degrade with time progression and their presence is governed by a set of rules (reactions). In a reaction, species can have the role of reactants, inhibitors or products. A reaction can take place only if all reactants are available and all inhibitors are not. Depending on the type of reaction, products levels are either increased or decreased. We have shown that properties of biological systems can be expressed in a suitable temporal logic and verified on the finite state space of the network. We have illustrated our framework in the modeling of blood glucose regulation.

We are currently investigating how to enrich reactions with response time, representing the required time for yielding products [8]. This poses new questions on how our model with time constraints could be compared to other existing time concepts for instance that in timed automata or that in stochastic models like in [13,11].

Finally, we have a prototype implemented with Snakes [19] and we plan to use Snoopy [14] and connected tools (Marcie [20]) to simulate and analyze CTL formulae.

**Acknowledgments.** We would like to thank Michel Malo and the anonymous reviewers for their comments and insightful suggestions. This work was supported by the French project ANR BLANC 0307 01 - SYNBIOTIC.

## References

1. R. Brijder, A. Ehrenfeucht, M. G. Main, and G. Rozenberg. A tour of reaction systems. *Journal of Foundations of Computer Science*, 22(7):1499–1517, 2011.

2. R. Brijder, A. Ehrenfeucht, and G. Rozenberg. Reaction systems with duration. In J. Kelemen and A. Kelemenová, editors, *Computation, Cooperation, and Life*, volume 6610 of *LNCS*, pages 191–202. Springer, 2011.
3. A. Cerone and A. Maggiolo-Schettini. Time-based expressivity of time petri nets for system specification. *TCS*, 216(1-2):1–53, 1999.
4. C. Chaouiya, A. Naldi, E. Remy, and D. Thieffry. Petri net representation of multi-valued logical regulatory graphs. *Natural Computing*, 10(2):727–750, 2011.
5. E. M. Clarke, O. Grumberg, and D. Peled. *Model checking*. MIT Press, 2001.
6. V. Danos and C. Laneve. Formal molecular biology. *TCS*, 325(1):69–110, 2004.
7. M. DeCristofaro and K. Daniels. Toxicogenomics in biomarker discovery. In D. Mendrick and W. Mattes, editors, *Essential Concepts in Toxicogenomics*, volume 460 of *Methods in Molecular Biology<sup>TM</sup>*, pages 185–194. Humana Press, 2008.
8. C. Di Giusto, H. Klaudel, and F. Delaplace. Reaction networks with delays applied to toxicity analysis. Technical report, IBISC, 2014. Available at .
9. E. Emerson and E. M. Clarke. Using branching time temporal logic to synthesize synchronization skeletons. *Science of Comp. Programming*, 2(3):241 – 266, 1982.
10. W. R. Foster, S.-J. Chen, A. He, A. Truong, V. Bhaskaran, D. M. Nelson, D. M. Dambach, L. D. Lehman-McKeeman, and B. D. Car. A retrospective analysis of toxicogenomics in the safety assessment of drug candidates. *Toxicologic pathology*, 35(5):621–35, Aug. 2007.
11. D. Gilbert, M. Heiner, F. Liu, and N. Saunders. Colouring space - a coloured framework for spatial modelling in systems biology. In J. M. Colom and J. Desel, editors, *Petri Nets*, volume 7927 of *Lecture Notes in Computer Science*, pages 230–249. Springer, 2013.
12. H.-M. Hanisch, K. Lautenbach, C. Simon, and J. Thieme. Timestamp petri nets in technical applications. In *WODES '98*, pages 321–326, 1998.
13. M. Heiner, D. Gilbert, and R. Donaldson. Petri nets for systems and synthetic biology. In M. Bernardo, P. Degano, and G. Zavattaro, editors, *SFM*, volume 5016 of *Lecture Notes in Computer Science*, pages 215–264. Springer, 2008.
14. M. Heiner, M. Herajy, F. Liu, C. Rohr, and M. Schwarick. Snoopy - a unifying petri net tool. In S. Haddad and L. Pomello, editors, *Petri Nets*, volume 7347 of *Lecture Notes in Computer Science*, pages 398–407. Springer, 2012.
15. K. Jensen. *Coloured Petri Nets - Basic Concepts, Analysis Methods and Practical Use - Volume 1*. EATCS Monographs on TCS. Springer, 1992.
16. J. Kleijn and M. Koutny. Membrane systems with qualitative evolution rules. *Fundam. Inform.*, 110(1-4):217–230, 2011.
17. J. Kleijn, M. Koutny, and G. Rozenberg. Modelling reaction systems with petri nets. In *BioPPN-2011*, pages 36–52, 2011.
18. A. Paun, M. Paun, A. Rodríguez-Patón, and M. Sidoroff. P systems with proteins on membranes: a survey. *International Journal of Foundations of Computer Science*, 22(1):39–53, 2011.
19. F. Pommereau. Quickly prototyping Petri nets tools with SNAKES. *Petri net newsletter*, (10-2008):1–18, 10 2008. SNAKES is available here.
20. M. Schwarick, M. Heiner, and C. Rohr. Marcie - model checking and reachability analysis done efficiently. In *QEST*, pages 91–100. IEEE Computer Society, 2011.
21. L. Serrano. Synthetic biology: promises and challenges. *Molecular Systems Biology*, 3(158), 2007.
22. C. B. Thanh, H. Klaudel, and F. Pommereau. Petri nets with causal time for system verification. *ENTCS*, 68(5):85–100, 2002.
23. M. D. Waters and J. M. Fostel. Toxicogenomics and systems toxicology: aims and prospects. *Nature reviews. Genetics*, 5(12):936–48, Dec. 2004.