PRES LUNAM
École Doctorale STIM
Sciences et Technologies de l'Information et Mathématiques

Spécialité : Informatique
Laboratoire : IRCCyN
Equipe : MeForBio

Presentation of the Process Hitting framework and inference of Biological Regulatory Networks with Thomas parameters

Folschette, Maxime
Mél : Maxime.Folschette@irccyn.ec-nantes.fr

Abstract: In this paper, the Process Hitting (PH), a recently introduced framework to model Biological Regulatory Networks (BRNs), is introduced. On the other hand, the qualitative modeling of BRNs has been widely addressed using René Thomas formalism, which is also depicted. A translation from PH to Thomas representation of BRNs relying on an exhaustive search of all regulations is finally presented.

Keywords: *qualitative modelling, model abstraction, interaction graph, René Thomas' parameters*
Collaborations : Loïc Paulevé, Katsumi Inoue, Morgan Magnin, Olivier Roux

1 Introduction

Biological Regulatory Networks (BRNs) consist in sets of either positive or negative mutual effects between some components (genes, proteins...). A discrete modeling approach was initiated by René Thomas in 1973, allowing the representation of the different levels of a component, such as concentration or expression levels, as integer values. Nevertheless, these dynamics can be precisely established only with regard to some "focal points", related to as Thomas' parameters, indicating the evolutionary tendency of each component.

In order to address the formal checking of dynamical properties within very large BRNs, we recently introduced in [1] a new formalism, named the *"Process Hitting"* (PH). A PH describes, in an atomic manner, the possible evolutions of a "process" (representing one component at one level) triggered by the hit of at most one other process in the system. This particular structure makes the formal analysis of BRNs with hundreds of components tractable [2].

In [3] we showed that starting from one PH model, it is possible to find the underlying interactions, then the underlying Thomas' parameters. This method relies on an exhaustive search of the interactions between components of the PH model, and an enumeration of the (possibly large) nesting set of valid parameters, so that the resulting dynamics are ensured to respect the PH dynamics.

2 Frameworks

2.1 The Process Hitting framework

A Process Hitting (PH) (Def. 1) gathers a finite number of concurrent *processes* grouped into a finite set of *sorts*. A sort stands for a component of the system while a process, which belongs to a unique sort, stands for one of its expression levels. A process is noted a_i where a is the sort and i is the process identifier within the sort a. At any time, exactly one process of each sort is present; a *state* of the PH corresponds to such a set of processes.

The concurrent interactions between processes are defined by a set of *actions*. Actions describe the replacement of a process by another of the same sort conditioned by the presence of at most one other process in the current state. An action is denoted by $a_i \to b_j \Rsh b_k$, which is read as "a_i *hits* b_j to make it bounce to b_k".

Definition 1 (Process Hitting) *A* Process Hitting *is a triple* $(\Sigma, L, \mathcal{H})$*, where:*

- $\Sigma = \{a, b, \dots\}$ *is the finite set of* sorts*;*
- $L = \prod_{a\in\Sigma} L_a$ *is the set of states with* $L_a = \{a_0, \dots, a_{l_a}\}$ *the finite set of* processes *of sort* $a \in \Sigma$ *and* l_a *a positive integer, with* $a \neq b \Rightarrow L_a \cap L_b = \emptyset$*;*
- $\mathcal{H} = \{a_i \to b_j \Rsh b_k \in L_a \times L_b \times L_b \mid (a, b) \in \Sigma^2 \wedge b_j \neq b_k \wedge a = b \Rightarrow a_i = b_j\}$ *is the finite set of* actions*.*

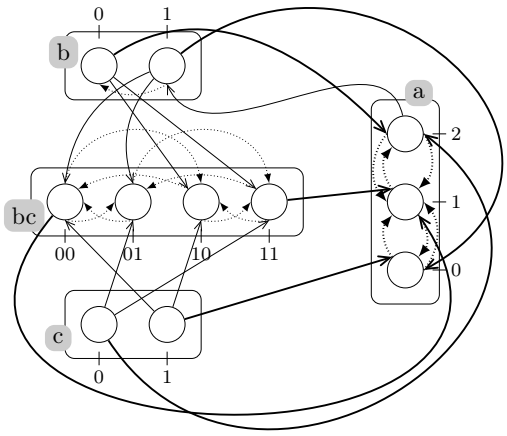


Figure 1: A PH example with four sorts: three components (a, b and c) and a cooperative sort (bc). Actions targeting processes of a are in thick lines.

Given a state $s \in L$, the process of sort $a \in \Sigma$ present in s is denoted by $s[a]$. An action $h = a_i \to b_j \Rsh b_k \in \mathcal{H}$ is *playable* in $s \in L$ if and only if $s[a] = a_i$ and $s[b] = b_j$. In such a case, $(s \cdot h)$ stands for the state resulting from the play of the action h in s, with $(s \cdot h)[b] = b_k$ and $\forall c \in \Sigma, c \neq b, (s \cdot h)[c] = s[c]$.

Modeling cooperation. As described in [1], the cooperation between processes to make another process bounce can be expressed in PH by building a *cooperative sort*. Fig. 1 shows an example of a cooperative sort bc between sorts b and c, defined with 4 processes (one for each sub-state of b and c). Each process of sorts b and c hit bc, to update its active process and, finally, the action $bc_{11} \to a_1 \Rsh a_2$ represents the cooperation between processes b_1 and c_1 to make a_1 bounce to a_2, instead of independent hits from b_1 and c_1.

Example 1 *Fig. 1 represents a PH* $(\Sigma, L, \mathcal{H})$ *with* $\Sigma = \{a, b, c, bc\}$*, and:*

$$L_a = \{a_0, a_1, a_2\}, \qquad L_b = \{b_0, b_1\}, \qquad L_c = \{c_0, c_1\}, \qquad L_{bc} = \{bc_{00}, bc_{01}, bc_{10}, bc_{11}\}.$$

This example models a BRN where components b and c are Boolean, component a has three qualitative levels and bc is a cooperative sort. In this BRN, a inhibits b at level 2 while b and c activate a with independent actions (e.g. $b_0 \to a_2 \Rsh a_1$*) or through the cooperative sort bc (e.g.* $bc_{11} \to a_1 \Rsh a_2$*).*

2.2 Thomas' modeling

Thomas' formalism, here inspired by [4, 5], first relies on an *Interaction Graph* (IG) to model the structure of the system by defining the components' mutual influences. The nodes of this graph represent components, while its edges stand for either positive or negative interactions (Def. 2).

Definition 2 (Interaction Graph) *An* Interaction Graph *(IG)* (Γ, E_+, E_-) *is a triple where:*

- Γ *is a finite number of* components*,*
- E_+ *(resp.* E_-*)* $\subset \{a \xrightarrow{t} b \mid a, b \in \Gamma \wedge t \in [1; l_a]\}$ *is the set of positive (resp. negative)* regulations *between two nodes, labeled with a* threshold *;* l_a *denotes the maximum level of component a.*

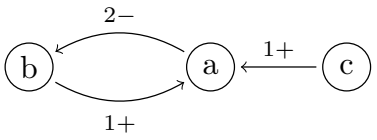
A regulation from a to b is unique, i.e. if $a \xrightarrow{t} b \in E_+$ *(resp.* E_-*), then there is no regulation* $a \xrightarrow{t'} b$ *in* E_- *(resp.* E_+*), and no other regulation* $a \xrightarrow{t''} b$ *in* E_+ *(resp.* E_-*) with* $t'' \neq t$*.*

For an interaction of the IG to take place, the expression level of its head component has to be higher than its threshold; otherwise, the opposite influence is expressed. For any component $a \in \Gamma$, $\Gamma^{-1}(a)$ is the set of its regulators:

$$\Gamma^{-1}(a) = \{b \in \Gamma \mid \exists b \xrightarrow{t} a \in E_+ \cup E_-\} \ .$$

A *state* s of an IG (Γ, E_+, E_-) is an element of $\prod_{a \in \Gamma}[0; l_a]$; $s[a]$ refers to the level of component a in s.

The specificity of Thomas' approach lies in the use of discrete *parameters* to represent focal levels. We chose to use intervals as focal levels as it offers more expressivity than single values (Def. 3).



$K_{a,\{b,c\},\emptyset} = [2;2]$ $\quad$ $K_{b,\{a\},\emptyset} = [0;1]$
$K_{a,\{b\},\{c\}} = [1;1]$ $\quad$ $K_{b,\emptyset,\{a\}} = [0;0]$
$K_{a,\{c\},\{b\}} = [1;1]$
$K_{a,\emptyset,\{b,c\}} = [0;0]$ $\quad$ $K_{c,\emptyset,\emptyset} = [0;1]$

Figure 2: (left) IG example. Regulations are represented by edges labeled with their sign and threshold. For instance, the edge from b to a is labeled "1+", which stands for: $b \xrightarrow{1} a \in E_+$. (right) One admissible parametrization of the left IG.

Definition 3 (Discrete parameter and Parametrization) *Let $x \in \Gamma$ be a given component and A (resp. B) $\subset \Gamma^{-1}(x)$ a set of its* activators *(resp.* inhibitors*), such that $A \cup B = \Gamma^{-1}(x)$ and $A \cap B = \emptyset$. The discrete* parameter *$K_{x,A,B} = [i;j]$ is a non-empty interval so that $0 \leq i \leq j \leq l_x$. With regard to the dynamics, x will tend towards $K_{x,A,B}$ in the states where its activators (resp. inhibitors) are the regulators in set A (resp. B). In the case where $x \in K_{x,A,B}$, x does not evolve. The complete map $K = (K_{x,A,B})_{x,A,B}$ of discrete parameters for an IG is called a* parametrization *of this IG.*

At last, dynamics are defined in BRN in a unitary and asynchronous way: from a given state s, a transition to another state s' is possible provided that only one component a will evolve of exactly one level towards $K_{a,A,B}$, where A (resp. B) is the set of activators (resp. inhibitors) of a in s, provided that $a \notin K_{a,A,B}$ in s.

Example 2 *Fig. 2(left) represents the Interaction Graph (Γ, E_+, E_-) with $E_+ = \{b \xrightarrow{1} a, c \xrightarrow{1} a\}$ and $E_- = \{a \xrightarrow{2} b\}$. Fig. 2(right) gives a possible parametrization of this IG. In this BRN, the following transitions are possible: $\langle a_0, b_1, c_1\rangle \to \langle a_1, b_1, c_1\rangle \to \langle a_2, b_1, c_1\rangle \to \langle a_2, b_0, c_1\rangle \to \langle a_1, b_0, c_1\rangle$.*

3 Biological Regulatory Network inference

This section focuses on the inference of a complete BRN with Thomas' parameters from a given PH.

3.1 Interaction Graph inference

In order to infer a BRN, one has to find the Interaction Graph (IG) first. This step assumes that the studied PH defines two types of sorts: the sorts corresponding to BRN components, which will appear in the IG, and the cooperative sorts, as defined in Subsect. 2.1.

Inferring global influences of a predecessor b on a component a requires to find "local influences" from this predecessor first, by considering a given state of the PH and changing only the active process of b. The aim is to compare the set of processes towards which the component a will evolve, for each active process of b, leaving the active process of all the other sorts unchanged. Indeed, if after increasing the level of b (i.e. after activating a higher process of b) we notice that a tends to reach a higher (resp. lower) level, we can then deduce that b activates (resp. inhibits) a in this selected state.

If all local influences of b on a are the same (activations or inhibitions), we can deduce that the global influence of b on a. The related threshold is the lowest level of b for which we can detect such an influence. Otherwise, if a behavior cannot be represented as a BRN, an unsigned edge with no threshold is inferred.

Example 3 *When this method is applied to the PH of Fig. 1, the IG given in Fig. 2(left) is inferred.*

3.2 Independent Parameters inference

This subsection presents a method to infer independent discrete parameters from a given PH, as given in [1]. We suppose in the following that the IG inference of the PH does not contain any unsigned edge, and in each sort, all processes activating (resp. inhibiting) another component share the same behavior. In this way, each parameter $K_{a,A,B}$ of a component a, with $A, B \subset \Gamma^{-1}(a)$, can be inferred by watching the behavior of a in any state compatible with the sets of activators and inhibitors A and B. The parameter $K_{a,A,B}$ is the set of stable processes of a that will eventually be active, given that this set is a non-empty interval; otherwise, the parameter cannot be inferred.

Example 4 *All parameters of the PH in Fig. 1 can be inferred, giving the parametrization of Fig. 2(right).*

3.3 Admissible parametrizations enumeration

The previous inference step may leave several parameters undetermined, due to behaviors impossible to represent in a BRN. If it is not possible to refine the PH model in order to remove these inconclusive cases, one can perform a last step to enumerate all valid values for all non-inferred parameters. We consider that a parameter is valid if any transition it involves in the resulting BRN is allowed by the studied PH dynamics. We also add some biological constraints on the whole parametrizations, given in [5]. These constraints lead to a family of admissible parametrizations.

3.4 Implementation

The inference method described in this paper has been implemented as a tool named `ph2thomas`, as part of PINT[1], a library gathering PH related tools. Our implementation mainly consists of Answer Set Programming (ASP) programs solved with Clingo[2]. ASP turned out to be effective for this work as it efficiently tackles the inherent complexity of the methods and the enumeration of admissible parametrizations.

Our approach can successfully handle large PH models of BRNs found in the literature such as a T-cells receptor model from [6] which contains 40 components[3], on which IG and parameters inferences are performed together in less than a second on a standard desktop computer.

4 Conclusion

This work establishes the abstraction relationship between PH, which is more abstract and allows incomplete knowledge on cooperations, and Thomas' approach for qualitative BRN modeling. This motivates the concretization of PH models into a set of compatible Thomas' models in order to benefit from the complementary advantages of these two formal frameworks and extract some global information about the influences between components.

As an extension of the present work, we plan to explore new semantics of BRNs to be able to tackle influences currently represented by unsigned edges.

Acknowledgment. This work was partially supported by the Fondation Centrale Initiatives.

References

[1] Loïc Paulevé, Morgan Magnin, and Olivier Roux. Refining dynamics of gene regulatory networks in a stochastic π-calculus framework. In *Transactions on Computational Systems Biology XIII*, pages 171–191. Springer, 2011.

[2] Loïc Paulevé, Morgan Magnin, and Olivier Roux. Static analysis of biological regulatory networks dynamics using abstract interpretation. *Mathematical Structures in Computer Science*, 22(04):651–685, 2012.

[3] Maxime Folschette, Loïc Paulevé, Katsumi Inoue, Morgan Magnin, and Olivier Roux. Concretizing the process hitting into biological regulatory networks. In David Gilbert and Monika Heiner, editors, *Computational Methods in Systems Biology*, Lecture Notes in Computer Science, pages 166–186. Springer Berlin Heidelberg, 2012.

[4] Adrien Richard, Jean-Paul Comet, and Gilles Bernot. *Modern Formal Methods and App.*, chapter Formal Methods for Modeling Biological Regulatory Networks, pages 83–122. 2006.

[5] Gilles Bernot, Franck Cassez, Jean-Paul Comet, Franck Delaplace, Céline Müller, and Olivier Roux. Semantics of biological regulatory networks. *Electronic Notes in Theoretical Computer Science*, 180(3):3 – 14, 2007.

[6] Steffen Klamt, Julio Saez-Rodriguez, Jonathan Lindquist, Luca Simeoni, and Ernst Gilles. A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics*, 7(1):56, 2006.

[1] Available at `http://process.hitting.free.fr`
[2] Available at `http://potassco.sourceforge.net`
[3] This model is available as an example distributed with PINT.