Review of Constructing Biological Regulatory Networks from Process Hitting models

July 24, 2014

General remark

The authors proposed a new formalism called Process Hitting to model dynamics of biological networks. In the article, they address the translation of PH into RT's models and Interaction graph. The process Hitting may provide a clean model for studying properties of biological network. The definition of PH is clearly explained p. 5. The work appears to be substantial containing theory, modeling and software development. However, some definitions remain really obscures (see detail of my comments) and the text gives the feel that the authors do not really pay attention to be understood in the formal part.

Beside, the definition of RT model, including signs that can be deduced from the K parameters must be justified or clarify. For a computer scientist it is always surprising to have redundant pieces of information without accounting their matching. With your definition I can write $a \xrightarrow{+,1} b$ and I can set $K_{b,\emptyset} = 1$ and $K_{b,\{a\}} = 0$. On the one side I define a as an activator and on the other side I define a as an inhibitor with K parameters. As signs do not enter in the interpretation, why they are included in the formalism as necessary "comments"?

In the current state of the article, except for the definition of PH, the formalism is hard to understand, sometimes insufficiently illustrated. PH theory is only developed and use by the team of the authors as far as I know. Therefore, I don't think that anyone can assimilate all the definitions but members of the team, by contrast to Petri net for example which is widely used in biological modeling. Thereby, to popularize this work and PH theory a supported pedagogic effort must be undertaken for explaining the motivations, giving the intuition of the methods (pp 14-15) and a particular

attention should be paid to illustrate the definition by examples. A guideline biological example is certainly welcome in this endeavor (e.g. λ phage).

Detail of comments

P. 1 Abstract: the following sentence is not understandable. Either more explanations are needed or simply suppress it. *"Its major originality lies in a specific restriction on the causality of actions, which makes the formal analysis of very large systems tractable."*

"On the other hand, the qualitative modeling of BRNs has been widely addressed using Ren Thomas formalism, leading to numerous theoretical work and practical tools to understand emerging behaviors."

The BRNs were introduced by S. Kauffman in 1969. His work also covers "numerous theoretical works". Therefore, it would be interesting to mention this fact. René Thomas is not the father of BRN applied to regulatory circuit dynamics. Beside, the definition of the BRN for S. Kaufman consists of a Boolean system dynamics with a Boolean function which is a normal way to define a dynamical system. The R.T.'s formalization is slightly unconventional, namely not based on a stepwise function.

While the formal checking of dynamical properties is often limited to small networks because of the state graph explosion, an other major drawback of this framework is the difficulty to specify Thomas parameters, especially for large networks requires more explanations on the condition of limitation. For Boolean networks, symbolic methods and parsimonious generation of states exist to circumvent the generation of the whole state graph and then scaling up the performances of the analysis (e.g. equilibria computation). Therefore, claiming that it is a "major difficulty" is partly true only. The following works are related to this topics:

- Corblin, F., Tripodi, S., Fanchon, E., Ropers, D., & Trilling, L. (2009). A declarative constraint-based method for analyzing discrete genetic regulatory networks. Bio Systems, 98(2), 91104.
- Bollman, D., Coln-Reyes, O., & Orozco, E. (2007). Fixed points in discrete models for regulatory genetic networks. EURASIP journal on bioinformatics & systems biology, 2007, 97356.
- Delaplace, F., Klaudel, H., Melliti, T., & Sen, S. (2012). Analysis of modular organisation of interaction networks based on asymptotic dynamics. In D. Gilbert & M. Heiner (Eds.), CMSB 2012

- Scalable Steady State Analysis of Boolean Biological Regulatory Networks Ay F, Xu F, Kahveci T (2009) Scalable Steady State Analysis of Boolean Biological Regulatory Networks. PLoS ONE 4(12): e7992.

P. 3 & 6 The term "cooperation" is often mentioned. It is not self-explanatory for the reader. In particular in P.3 "associated influences" is not explanatory at all. If I have missed it I suggest to more clearly emphasize it. This comment does not concern cooperative sort which is clearly explained, but the notion of cooperation that cooperative sort encompasses. A definition of "cooperation" is required (formal or not) or more explanations of the notion of cooperation in this context.

P. 6 The arrows representing the hitting of sort a fig 1. are too small.

P. 8 *This allows the existence of interleaving of actions leading .. semantics with prioritized actions defined in [13].* Again I don't understand the message related to your own notion of cooperation. Is it really important to mention these elements?

P. 9 Interaction Graph. Γ refers to the set of vertices. However $\Gamma^{-1}(b)$ requires to seek the pre-image in E. I find the choice of Γ to annotate a function giving the regulators confusing. If we admit that E defines a relation, Eb or $E^{-1}(b)$ are more natural notations for the pre-image or more simply $b \rightarrow$ for the image and $\rightarrow b$ for the pre image. These notations are classically used for relation to qualify these sets of elements and do not require additional notations.

P. 8 You mention that *Thomas formalism lies on two complementary descriptions of the system* IG and K parameters. However There exists some redundancies for these description because the sign may be deduced from the parameters representing monotonous properties of the parameters. Moreover, The dynamics is not deduced from signs. IG graph with levels and parameters appears to suffice for fully describing the RT model. It would be more parsimonious (for an article in TCS) to define an IG with levels only, and parameters corresponding to Thomas' model and then explain that signs are abstractions of the parameters corresponding to monotony properties.

P. 11 Definition of the sign. The term of monotony is classically associated to the inference of signs and should be clearly mentioned. Here it is hidden in your definition. The discussion on the number of maximum levels has already been fully developed in other articles related to RT formalisms. Is it necessary here just for mentioning that PH may accept any level?

P. 12 The order of your explanations makes the text confusing. Signs are components of the definition of an IG. After you argue that it is useless because parameters suffice to deduce the state graph. You mention that they can be inferred although they must be included in the definition of an IG. The formalization is not clear. Unsigned IG could/must be presented before signed IG which is an abstraction of the monotony of the K parameters.

P. 12 *However, it is common that the prior IG actually refers to interaction .. enhancing the conclusiveness of static analyses* The inference of an IG from a dynamics is interesting and important in biology. However the motivations expressed here are not well argued. In particular I don't find this "common". Some bibliographical references are required to consolidate the arguments or the sentence has to be rephrased.

P. 13 -14 - 15 3 pages contain 5 definitions with no example and no intuition on their use but for Γ and focal. More explanations are required: what kinds of concept do represent focal, what is the importance of a strict context, where is the outline guiding the reader through the dense area of definitions located in three pages only? These pages must be rewritten and extended to understand the core of the method.

P. 18 Figure 5 shows a result of Fig 3. However the detail concerns example in Fig. 1?

Examples and the implementation part are more clear. However, the biological example is only used to evaluate performances of the implementation. A concrete (biological) modeling example used as a guideline for illustrating the different notions throughout the article would certainly help the reader.