Journal of Theoretical Biology 403 (2016) 38–44

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Theoretical knock-outs on biological networks

Pedro J. Miranda [a,*], Sandro E. de S. Pinto [a], Murilo S. Baptista [b], Giuliano G. La Guardia [c]

[a] Department of Physics, State University of Ponta Grossa, Paraná, Brazil
[b] Institute for Complex System and Mathematical Biology, SUPA, University of Aberdeen, Aberdeen, United Kingdom
[c] Department of Mathematics and Statistics, State University of Ponta Grossa, Paraná, Brazil

HIGHLIGHTS

- A generic model applicable to biological phenomena described by directed graphs.
- A way to generate the order of importance of biological agents.
- A random walk model for a directed graph based on biological phenomena.
- Two methods to compute the flux of walkers in a directed graph.
- A diffusion model of "stimuli" in a biological network.

ARTICLE INFO

Article history:
Received 1 December 2015
Received in revised form
28 March 2016
Accepted 12 May 2016
Available online 14 May 2016

Keywords:
Relational biology
(M, R)-system
Complex networks
Random walks
Theoretical KOs

ABSTRACT

In this work we redefine the concept of biological importance and how to compute it, based on a model of complex networks and random walk. We call this new procedure, theoretical knock-out (KO). The proposed method generalizes the procedure presented in a recent study about Oral Tolerance. To devise this method, we make two approaches: algebraically and algorithmically. In both cases we compute a vector on an asymptotic state, called flux vector. The flux is given by a random walk on a directed graph that represents a biological phenomenon. This vector gives us the information about the relative flux of walkers on a vertex which represents a biological agent. With two vector of this kind, we can calculate the relative mean error between them by averaging over its coefficients. This quantity allows us to assess the degree of importance of each vertex of a complex network that evolves in time and has experimental background. We find out that this procedure can be applied in any sort of biological phenomena in which we can know the role and interrelationships of its agents. These results also provide experimental biologists to predict the order of importance of biological agents on a mounted complex network.

1. Introduction

Biological phenomena often rely on a variety of complex dependencies, feed-backs, and auto regulations. As we try to build a general theory to understand these phenomena, one may find it difficult to uncover a common ground to dissertate about biological central questions. One of the most influential theoretical biologists was Nicholas Rachevsky. He introduced the idea that those phenomena could be approached into two different ways: the relational and metrical aspects of biological systems (Rashevsky, 1948).

The relational biology deals with the complexity and relationships observed on well defined biological agents, such as: enzymes, cells, tissues, organs, species, etc. This sort of approach leads to general structures which can mainly be modeled by graphs, in the terms of graph theory, which is consistent on how reliant on complex structures biological phenomena are. Most of the complex network theory (*i.e.*, graphs that models real entities) developed today for modeling biological systems is a remnant of Rachevsky view point of biological processes. However, there are few works that recognize this author's contribution to this field, and we find a wealth of concepts that can be availed in order to understand biological processes.

On the other hand, there is the metrical biology that encompasses the reducionistic approach of modern biology and biotechnology. This approach allows one to know the biological structures involved on a phenomenon in details, but it rely mostly on the assumption that structure imply function (Cottam et al., 2007). In other words, the knowledge of smalls structures involved

* Corresponding author.
E-mail address: pedrojemiranda@hotmail.com (P.J. Miranda).

http://dx.doi.org/10.1016/j.jtbi.2016.05.021
0022-5193/

in a phenomenon implies in the knowledge of all its function and leads ultimately to the understanding, description, and prediction of the subject matter.

We recognize that modern biology has shed light on much of what we actually know about behavior of biological systems by the knowledge of its parts. However, there still exist difficulties which pervade general realizations of biology itself. Some of the main questions of biology can exemplify these difficulties: what is preserved in life phenomena; how structures are related to give rise to life; how life behaves in time and space, and so on. In order to answer and reconcile these questions it is necessary to build a general theory that turns it possible to approach biology conceptually and experimentally. Modern biology have shown that it has no discourse to offer this theory, since biological phenomena are resilient to final reduction in terms of Cartesian Method (Cottam et al., 2007). This fact of resilience is due to the nature of these phenomena and how we understand them. The lack of a general theory into a metrical approach for biology, pull theoretical biologists towards the relational biology language. One of the most important contribution to this area was the model introduced by Rosen (1958). In his work, Rosen proposed a metabolic network's model that takes into account three metabolycal processes: anabolism, catabolism, and repair. This model is known as $(\boldsymbol{M}, \boldsymbol{R})$–*system*, and it is composed by a set of *components* $\boldsymbol{M_i}$ of the system $\boldsymbol{M}$. This system is a connected directed graph in which the vertices are components (*i.e.,* representatives of biological agents) and the directed edges are input and outputs to the components.

The inputs are directed edges that points toward a component and outputs are directed edges that points from components. We interpret inputs and outputs as materials to be utilized by components in order to generate outputs. In this theory, Rosen differentiate coarse structures and fine structure. The latter relates to abstract systems in which vertices are "black boxes", which it is known only its input materials and output materials, but not how it operates; and the former relates to specific known systems of cells, enzymes, tissues, organs – which are results of metrical biology that convey the bridge between the relational biology (Rosen, 1958). After this work, Rosen realized that is wasnot enough to describe most metabolycal phenomena and used this introductory work as a background to posterior formalization: the categorification of $(\boldsymbol{M}, \boldsymbol{R})$–*system via* category theory (Rosen, 1958, 1959).

A good general theory for a field like biology should allows one to test experimentally, or numerically, each step that the theory when it is developed and detailed. By this assumption we mean that the most propositions should be testable to turn the theory intelligible and scientifically valid. Studying Rosen's model, we find an important proposition that took our attention: the importance of each component of a $(\boldsymbol{M}, \boldsymbol{R})$–*system*. Rosen defend that certain components are more important in the operation of the system, and take as a measure of the importance the *number of environmental outputs* of the system that cease to be produced due to a component inhibition (Rosen, 1958). This informal definition of a vertex (*i.e.,* component) importance in a biological systems accounts only on the system outputs to the environment, and has little to do with the internal implication of a vertex inhibition.

If we consider a *non-central component* of a $(\boldsymbol{M}, \boldsymbol{R})$–*system*, which is a vertex that if removed would not result in the failure of the entire system, this vertex still should cause more damage than imagined since the mutual reliance of vertex is frequently found in biological systems. For example, the failure of a secondary cytokine can generate debilitation on the complex network in which that cytokine play a role. This is observable because other cytokine, being primary, should compensate the lack of the former inhibited cytokine and the dynamics of the network changes internally, but is generating the same sort and amount of environment outputs. However, when time goes on, the systems changes completely by the inhibition of that secondary cytokine, and some pathology or system malfunction may arise. In another example, lets us consider a cell where exists many enzymes which have the same substrate. The inhibition of them causes a demand in the other, altering the quantity of products generated by the same substrate. As in the case of cytokines, the inhibition of a secondary enzyme causes the primary to be overused, and more transcription of the gene that corresponds that enzyme must be performed. As time progresses, the cellular network will change internally while it is trying to keep its outputs in an acceptable level. A species that have a defined niche, when inhibited in an ecological system, shall cause less competition on species of same niche and limited resources. This inhibition cause variation on how other species are related ultimately changing the structure and dynamical dependences of the complex system it is modeled by. Still, a view from "outside" of the ecological system does not demonstrate great differenced before and after removing a species (unless it is a central species), since most environmental outputs are being produced.

Besides subtle, these examples illustrate how deeply is the consequence of removing a system's component. It also illustrates how the alteration of the system's outputs, by itself, should not suffice to quantify the degree of importance of a component. In order to remove the subtlety of this conception, this work has as aim to generate a steady quantification of the importance of each component involved in a biological network. Inspired on knockout (KO) on animal models, in which a gene is suppressed and a population of animal subjects is knocked-out for its corresponding phenotype, we introduce the concept of theoretical KO, which is the effect of the removal of a vertex and how it affects the biological network in which it belongs.

We also want to emphasize that such importance quantification changes the previous concept of biological importance given by Rosen's work. We defend that this concept is based on the participation of the biological agents in relation to the whole network that contains it. In other words, the importance of biological agents must rely on how they change the internal dynamics when putted away. We devised this conceptual background from the generalization of a previous study about Oral Tolerance (Miranda et al., 2015). We noted that such method could be applicable in any sort of biological network that the notion of biological importance is ubiquitous. In other words, the main contribution of this work is to redefine biological importance in a complex background and to quantify such importance by utilizing a standard procedure.

We can apply this procedure to any biological phenomena that can be modeled by a connected directed graph in which dynamical processes can be used. We proceed with this endeavor using random walks in directed graph model, exploring both analytical and computational of this stochastic dynamics. To organize our specific goals to achieve this particular objective we list:

i. propose a generic model in which it is possible to use random walk in directed graphs to define and to calculate an invariant quantity that allows to measure the importance of any vertices of a graph (*i.e.,* complex network);
ii. propose an analytical method to find the same invariant quantity due to the random walk in complex networks, and when applicable;
iii. propose an algorithm that encompasses the sequence of operation over the network to generate statistically the defined invariant quantity;
iv. defend and discuss the experimental utility of the model for predicting important theoretical KOs.