

Comparative Netomics - lessons from cross-disciplinary network comparison (tentative title)

Throughout the history of science, we have seen many examples in which the advancements of biology have been catalyzed by discoveries in other disciplines. For instance, the maturation of X-ray diffraction facilitated the discovery of double helix, and later on the characterization of structures of thousands of different proteins. In the era of systems biology, attentions have been shifted from individual molecular components to their interactions in a system level. One may wonder from which discipline will biology be benefited most. In this essay, we want to argue that, while the influx of ideas in the past was most originated from specific areas in physics or chemistry, the new wave of catalysts come from areas of science very far apart, as different as engineering, behavioral science, sociology, but centered on the concept of network.

Network is by no mean new to biologists. Over the last decade, plenty of intertwined wiring diagrams were published in high profile journals. While many of us have been astonished by the complexity of such networks, few are able to gain any intuition from the hairballs. In this essay, we want to argue that, intuitions as well as mathematical methods developed in commonplace networks from other disciplines are able to catalyze our understanding of biology. Indeed, not only capturing the unique flavor of systems biology, the concept of network essentially describes the interactions between individual constituents-who is interacting with whom-in any complex system. The simple description thus enables one to examine the underlying skeletons of cross-disciplinary complex systems in the same ground as the interactions between molecular components in cells. The common network language makes ideas and methods developed to understand the organization of complex systems in diverse fields more accessible to biologists who want to gain insights on the complex interactions between molecular components, and vice versa. To leverage the interdisciplinary connections, it is instructive to initiate comparison between biological networks with networks in other disciplines.

Though underlying networks of various systems may resemble one another, comparing a bio-molecular network with a complex network from a distant field, say a social network, sounds comparing apples to oranges. What kinds of comparison could truly deepen our understanding? We believe there are a few points one should pay attention. First of all, the term "biological network" is used rather loosely in literature. It refers to all networks originated from any subfields of biology, including ecological networks or food webs. Here, however, we focus our attention to molecular networks coming from genomics or systems biology because it is in general harder to gain intuitions in such networks. Second, networks of particular interests to genomics or systems biology can be further be divided into two classes. The first class consists of examples like transcriptional regulatory networks, metabolic networks or protein-protein interaction networks. These examples are intrinsically different, for instance, a regulatory network describes part of the cellular information processing, a metabolic network traces the chemistry of metabolites, and the protein-protein interaction network is essentially a manual on how to assemble molecular machines. While these networks capture different facets of the complex organization of an organism, they all refer to specific goals of performance mediated by actual mechanistic interactions. Such networks closely resemble, and should be compared with networks with similar goals like networks from engineering or technological systems. There is another class of biological networks with examples like genetic interaction networks or disease networks. The former example is a sophisticated metrics summarizing results of double knockout experiments whereas the later describes the statistical association between genes and diseases as a result of GWAS and analysis of various genomics variants. Such networks do not show mechanistic interactions but a mathematical abstraction of complex relationships. From an algorithmic or method development standpoint, they share common graphical structures arise in computational social science, for instance mathematical machinery used in the bipartite network between genes and diseases can resonate with movie recommendation scheme building on a similar bipartite network between users and movies.

The striking similarity between biological systems and technological systems has long been identified. For instance, transcription regulation uses cooperative binding to arrive at a sigmoidal response curve whereas the same is achieved by an amplifier in circuit design. Analogies of this kind have inspired the advancement of synthetic biology. In terms of global network organization, a decade ago, Uri Alon pointed out several common design principles such as modularity and the usage of recurring elements (motifs) \cite{Alon Sci 2003}. These common topological patterns are in many cases the reflection of common underlying design objectives or strategies. For examples, modularity makes a system more evolvable, which is an advantage one would like to optimize. Therefore, an insightful network comparison should be rooted in the common design objectives. As Alon highlighted by the phrase “the tinkerer as an engineer”, it is remarkable that “good-engineering solutions” are found in biological systems evolved by random tinkering. Indeed, comparison between biological and technological networks should manifest the nature of the two very different approaches: evolution as a tinkerer trying to connect random nodes until the network is good enough to work, whereas technological networks are essentially blueprints drawn by engineers. Biologists often tend to distinguish the two approaches cautiously so as to avoid the notion of intelligent design – the existence of an intelligent cause that construct living organisms on purpose. Nevertheless, the distinction is not clear-cut. To a certain extent, an engineer is a tinkerer. Both biological networks and man-made technological networks like roadways and circuits are complex adaptive systems, there are plenty of examples showing that many great innovations are results of trial and error, and all technological systems are subjected to selection like users requirements. Tinkerer and engineer are fundamentally similar that in a recent review, Andreas Wagner summarized nine commonalities between biological and technological innovation, such as descent with modification, extinction and replacement, and horizontal transfer \cite{Wagner 2014}. In such a common framework, biological networks and networks with specific performance goals in other disciplines are shaped by similar underlying design objectives. However, as “there’s no free lunch”, there is no way to satisfy all objectives and thus tradeoffs are unavoidable. Networks in different disciplines are perhaps the optimal configurations subjected to different constraints and tradeoffs. Here, via a few specific examples, we want to illustrate how network comparison works in the broadened context. (~1000 words)

[[KKY: here a few paragraphs that describe specific examples of comparisons really bring intuition to biology which was poorly understood. Possible examples:

- In terms of information propagation/hierarchy, the comparison of regulatory network, management hierarchy
- Linux as an example to illustrate the bio vs tech comparison. Emphasize on the evolution versus tinkerer,
-

~600 words]]

[[KKY: a transition sentence, then a few more paragraphs describing examples on the share of common math. structures, and how methods developed in other disciplines are applied to biology. a reasonable way to divide the contents

- Algorithms deal with nodes, e.g. common guilt by association techniques used in bio. and social network, or PageRank for prioritizing importance, network stratification..
- Algorithms deal with links, e.g. link prediction
- Algorithms deal with organization. e.g. OrthoClust and many other modules detection scheme
- Methods deal with dynamics: control theory [[DW3, PM2]], circuit theory [[DW4]], information theory [[DW5]], game theory [[DW7]].

The big data story could be mentioned in these few paragraphs. Most of the examples we found have already been used by others in biological context, are there more novel suggestions? ~800-1000 words]]

[[KKY: the essay should be ended with an outlook, it will be great if there are a few concrete lessons, ~100 words]]