Heuristic Strategies in Systems Biology

Fridolin Gross[†]
fridolin.gross@uni-kassel.de

ABSTRACT

Systems biology is sometimes presented as providing a superior approach to the problem of biological complexity. Its use of 'unbiased' methods and formal quantitative tools might lead to the impression that the human factor is effectively eliminated. However, a closer look reveals that this impression is misguided. Systems biologists cannot simply assemble molecular information and compute biological behavior. Instead, systems biology's main contribution is to accelerate the discovery of mechanisms by applying models as heuristic tools. These models rely on a variety of idealizing and simplifying assumptions in order to be efficient for this purpose. The strategies of systems biologists are similar to those of experimentalists in that they attempt to reduce the complexity of the discovery process. Analyzing and comparing these strategies, or 'heuristics', reveals the importance of the human factor in computational approaches and helps to situate systems biology within the epistemic landscape of the life sciences.

keywords: systems biology, heuristics, computational modeling.

1. Introduction

There exists a widespread narrative about the relationship of systems biology to the traditional experimental approaches of molecular biology and cell biology. According to this narrative, the traditional approaches have followed an overly "reductionistic" strategy by looking at individual parts in isolation, whereas systems biology succeeds in assembling information about all relevant molecular players and translating this information into a "holistic" understanding of the system. This success is expected from the application of computational methods to integrate large and heterogeneous data sets and to simulate biological behavior:

[†] University of Kassel, Germany.

> Perhaps the most important consequence of the Human Genome Project is that it is pushing scientists toward a new view of biology—what we call the systems approach. Systems biology does not investigate individual genes or proteins one at a time, as has been the highly successful mode of biology for the past 30 years. Rather, it investigates the behavior and relationships of all of the elements in a particular biological system while it is functioning. These data can then be integrated, graphically displayed, and ultimately modeled computationally. (Ideker et al., 2001, p. 343)

The implicit criticism of the traditional approach is not simply that it is slow or inefficient, but that it is conceptually misguided: It underestimates the complexity of living systems and overlooks relevant phenomena by investigating biological processes at the wrong level.

> Biological systems are extremely complex and have emergent properties that cannot be explained, or even predicted, by studying their individual parts. The reductionist approach—although successful in the early days of molecular biology—underestimates this complexity and therefore has an increasingly detrimental influence on many areas of biomedical research, including drug discovery and vaccine development. (van Regenmortel, 2004, p. 1016)

Systems biology is presented as superior and less biased for several reasons: First, it takes into account all parts of the system instead of focusing on individual components. Second, it applies sophisticated statistical methods to analyze and integrate data instead of relying on subjective judgement. And third, it uses formal and quantitative models to simulate biological processes instead of the intuitive reasoning strategies of experimentalists. All of this suggests that systems biology manages to effectively eliminate the human factor from biological research.

In this contribution I want to develop a more balanced perspective on the relationship between systems biology and the traditional approach. I argue that systems biology does indeed manage to overcome some of the latter's biases, but it can develop productive research strategies only by itself introducing potential biases. Computational methods in biology will not in the foreseeable future gain the character of the robust, standardized, and algorithmic methods that may be found in some areas of engineering. Instead, they are only as good

as the assumptions that they are based on and thus dependent on our current limitations in scientific knowledge and cognitive capacity.

2. Scientific Discovery as Problem Solving

In order to analyze different research strategies in biology, I will start by introducing the concept of "heuristics." I understand heuristics, in line with scholars such as Herbert Simon and William Wimsatt, as problem-solving strategies that work by reducing the complexity of a given research task (e.g. Simon, 1962; Wimsatt, 2007). Within this framework scientific discovery is understood as a search through a problem space. This space is determined by the structure of the research problem and by concepts and parameters specifying possible solutions that usually stem from background theories and beliefs (Resnik, 1997). If the problem space is small, one may consider to use random search to test all candidate solutions, but in more complex situations this is not an efficient strategy. Heuristics are rules of thumb that facilitate the discovery process by restricting or directing the search through the problem space. They introduce specific assumptions about the research problem that may or may not be justified. Consequently, what makes these strategies efficient at the same time creates the risk of underestimating the complexity of the problem. Heuristics make the search selective, raising its efficiency over blind trial-and-error search, at the cost of introducing bias.

In biology a typical research task consists in explaining a phenomenon by identifying the underlying causal factors and by showing how these factors interact to bring about the phenomenon. There are thus two different aspects that potentially contribute to the complexity of this scientific task: determining the causal structure and figuring out how it works. Heuristic strategies can be used to address both of these aspects.

Heuristics of search work by restricting the set of possible components, interactions, or modes of organization, thereby guiding the discovery of the underlying causal structure. But even if the structure and internal organization of a system are fully known, it can be difficult to understand its behavior due to its intrinsic complexity. Only in very few situations there is a straightforward algorithm that can be applied to 'solve' it. In most cases one has to approach understanding via intermediate steps that transform the initial problem into a more manageable one. Here, heuristic strategies often work by introducing idealizations. The roles of intrinsic complexity and idealization gain

importance in systems biology. In traditional experimental biology, by contrast, the emphasis is on heuristics of search. However, the distinction between the two kinds of heuristics is not as clear-cut, nor is their assignment to different scientific approaches. Indeed, as I will argue in the following section, large parts of systems biology can be understood as contributing to the problem of search by exploiting the power of computational models as heuristic tools (see also MacLeod, 2015).

3. Heuristic Strategies in Experimental and Computational Approaches

According to the systems biologists I quoted in the introduction, traditional molecular biology has been useful in determining the parts of living systems, but an entirely different approach is needed when it comes to putting those parts back together in order to understand systemic behavior. This view presupposes that our current knowledge about the causal structures underlying biological phenomena is sufficiently reliable and complete. However, in spite of increasing amounts of data accumulated at the level of DNA, RNA and protein, of concentrations of metabolites, of epigenetic modifications etc., most areas in molecular biology are lacking knowledge about the causally relevant factors. At the same time many of the measurements, despite being quantitative, often lack in both precision and accuracy. Moreover, the measurements that are available from experiments are rarely of the kind that can immediately be used for simulating the dynamics of a system in a computational model. Accordingly, rather than trying to understand the behavior of systems whose internal structures are largely known, what many systems biologists do in practice is to use models to test hypotheses about underlying structure. The following quote nicely makes this point:

> We believe that modeling these important biological systems cannot wait until all the rates are reliably measured, or even until all the various players and interactions are discovered. Indeed, the most important role of modeling is to identify missing pieces of the puzzle. It is as useful to falsify models—identifying which features of the observed behavior cannot be explained by the experimentalists' current interaction network—as it is to successfully reproduce known results. (Brown et al., 2004, p. 185)

One of the main roles of computational models in systems biology is thus to facilitate the discovery of mechanisms. Modeling can be used as an additional tool to restrict the set of possible causal structures underlying a particular phenomenon.

3.1. The traditional approach

Before taking a closer look at how heuristic strategies are applied in systems biology, I will briefly characterize the research strategies of the traditional experimental approach. Obviously, there would be a lot to say about strategies of experimental design and the material aspects of scientific practice in general, but for the sake of comparison I will focus on cognitive strategies of experimental approaches in molecular and cell biology and do not claim to be exhaustive in any way. As I have argued in detail elsewhere, the traditional experimental approach relies on some general heuristic strategies that are characteristic of mechanistic science in general, but also on more specific ones that are rooted in the historical development of molecular biology (Gross, 2013). Among the more general strategies is the heuristic of *decomposition and localization* (Bechtel & Richardson, 1993) that simplifies the task of understanding a complex phenomenon by conceiving it as produced by simpler activities that are carried out by structurally defined components of the system. Biological systems are thus approached as consisting of quasi-independent modules, which obviously relies on strong assumptions about biological organization. Furthermore, experimentalists typically assume that the organization of a mechanism is relatively simple. Mechanisms are expected to consist of a manageable number of components and to be organized in a largely sequential fashion. William Bechtel has argued that the expectation of sequential organization is fundamentally linked to the human cognitive setup:

> The assumption of sequential order reflects the practices of many scientists, who attempt to envisage sequentially the qualitative changes occurring in the mechanisms they investigate. More fundamentally, this reflects the sequential nature of human mental processes. We perceive successive states of the world, and in imagination we redeploy perceptual processes (...) and so imagine changes sequentially. (Bechtel, 2011, p. 536)

But experimental biologists focus on sequential organization also because it resonates with a general perspective of biological processes as processes of information transfer. Here I do not primarily have in mind the conception of encoded information that is transferred from DNA to protein during the process of gene expression, but rather the more general idea of signals being transferred by a chain of molecular events.

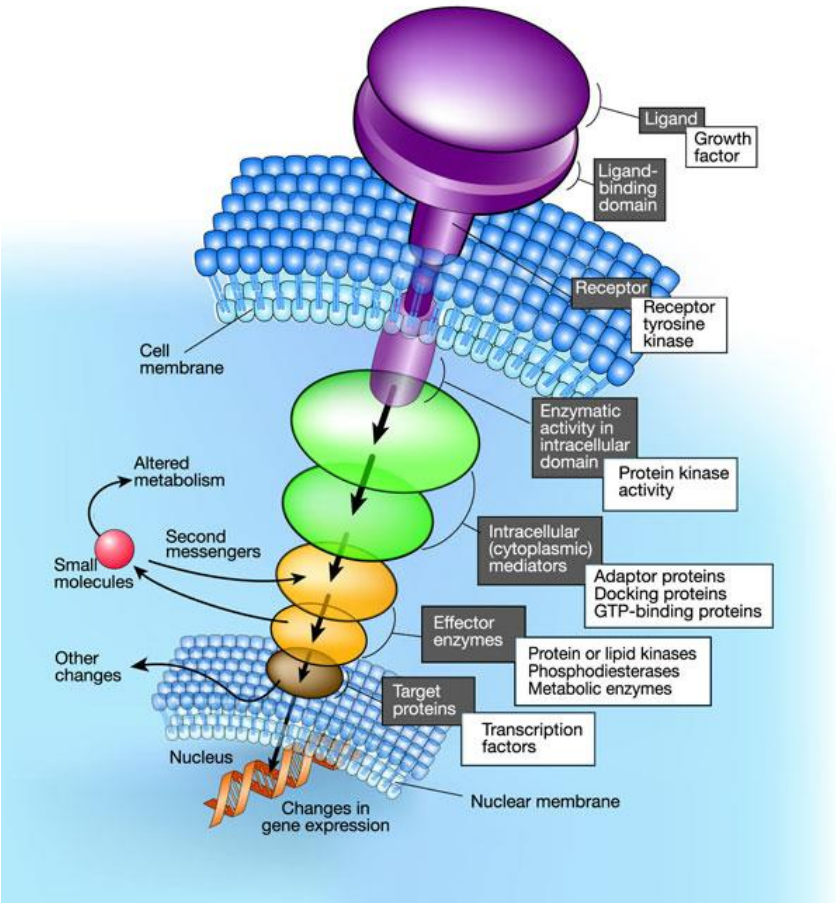

Figure 1: Schematic representation of a signal transduction pathway. *Source*: Downward (2001).

Figure 1 shows a particularly clear example of this, representing how an extracellular signal is transmitted to the nucleus in what biologists call 'signal transduction.' Apart from illustrating sequential organization, two additional aspects become clear from the style of representation of this diagram. First, even though it represents a chain of chemical reactions, it does not spell out in detail how these reactions proceed and influence each other. The explanatory burden lies in establishing the identity and order of the links in the chain. Principles of biochemistry may be invoked for the explanation of individual reaction steps, phosphorylation, inhibition, etc., but not for the dynamics of the overall process. The assumption that the transmission of the signal is largely unrestricted by chemical principles allows biologists to investigate the individual links in the sequence independently from each other. A second observation is that in order to explain the process, it is sufficient to represent it as if it were build up of interactions of individual molecules. Even though scientists are aware that the number of molecules engaging in each of the steps represented in the diagram is potentially very large, this kind of information is not needed in order to understand the mechanism. In general, they do not expect any non-trivial population effects in molecular processes. In summary, there are several specific assumptions that experimental biologists make about the expected types of organization which together significantly reduce the complexity of their research task and, in particular, allow them to investigate and explain biological phenomena largely without having to apply quantitative or computational methods.

This general qualitative strategy has been quite successful for explaining a wide range of biological phenomena, but as increasing amounts of information about molecular processes are collected, biologists are getting a clearer idea of the contexts in which the underlying assumptions may not be justified. In these contexts alternative strategies may be of help in the project of discovering mechanisms in order to explain biological phenomena.

3.2. Strategies of systems biology

How do strategies in systems biology relate to and interact with traditional experimental research? Instead of analyzing detailed case studies, I will confine myself here to some general characteristics. I will discuss three different ways in which systems biologists address the complexity of biological processes: small models, large models, and network approaches. While the first two are

usually subsumed under the label of ′bottom up,′ the third might be considered a case of ′top down modeling.′

3.2.1. Small models

Systems biologists sometimes use computational approaches to investigate the same mechanisms that are also studied by experimentalists. That is, they focus on an isolated phenomenon and create a mathematical model that comprises only those components that are thought to be relevant for the phenomenon. The motivation to build and study such a model often lies in the observation that the system shows puzzling or otherwise interesting dynamic behavior, such as bistability or oscillations. The resulting models are relatively coarse-grained and do not attempt to incorporate all known molecular details. Small models are used when the knowledge about the mechanism is incomplete and not much quantitative data are available. The aim is to achieve a better understanding of the dynamics of the mechanism and to show that it is possible to reproduce its qualitative behavior based on the most relevant causal interactions. A failure to reproduce general behavior points to important pieces that are missing in the description of the mechanism. Examples are models of the core mechanism of the circadian clock, of individual stages of the cell cycle, or of small signaling systems such as the p53 or NF-κB networks.

How can we understand the approach of building small models within the general framework of heuristics? First of all, this approach is based on the strategy of decomposition and localization that is also used by experimentalists because it focuses on a specific phenomenon and aims at an explanation in terms of the activities and interactions of individual components. However, the underlying mechanism is not described verbally or in a cartoon model, but with a formal mathematical language. This allows systems biologists to deal with more sophisticated and non-sequential forms of organization, such as feedback loops or cross-talk between pathways. Using a model, they achieve a ′recomposition′ of the mechanism that can explain aspects of behavior that may be hidden in the qualitative accounts of experimental biologists. Moreover, a quantitative framework can describe the kinetics of molecular processes, sometimes also their distribution in space, and thereby go beyond the purely qualitative understanding of a mechanism as a chain of discrete signaling steps. In this way, inconsistencies in biologists′ intuitive reasoning about mechanisms can be uncovered. In summary, building small models

allows systems biologists to overcome some of the more specific biases of the traditional approach. The power of this approach relies on the extended cognitive abilities provided by computational modeling.

But an increase in power does not necessarily mean that bias is reduced: models always introduce assumptions of their own. The construction of a model is never fully determined by data or underlying theories alone, but relies on a number of often very pragmatic choices (see Morgan and Morrison 1999). This is particularly obvious in the case of small models. First, as already mentioned, these models usually skip a lot of molecular detail. This is done partly because more specific information is simply lacking, but also with the aim to restrict the number of undetermined parameters. If a model contains more free parameters than can be determined by available data, the modeling problem is not well-constrained and the results based on this model will not necessarily be meaningful. Therefore, components of small models can often not be directly assigned to particular molecular players, but lump together the net influence of several causal factors. Second, while modelers make use of underlying biochemical principles, they commonly do so in the form of highly idealized reaction schemes, such as the law of mass action or Michaelis-Menten kinetics. Furthermore, most models are expressed in the form of ordinary differential equations (ODEs) and therefore neglect spatial effects by assuming that all reacting molecules are well-mixed. All these idealizations and assumptions are introduced in order to keep the model manageable and useful as a heuristic tool. In this context, systems biologists usually do not merely aim at computational tractability, but also demand that the simulations of a model can be followed and eventually comprehended, even if the results are counterintuitive. Thus the ′human factor′ substantially shapes the modeling strategy, even if some of the cognitive procedures are replaced or extended by algorithmic procedures.

Systems biologists directly build on the findings of molecular biology, and they are interested in the solution of the same epistemic puzzle of how a mechanism works and produces the phenomenon of interest. Yet, they propose a modified strategy for solving this puzzle by introducing computational methods and going beyond the qualitative schemes proposed by experimental biologists. Describing a phenomenon in quantitative terms and hypothetically recomposing the underlying causal structure while taking into account specific physical and biochemical constraints, allows systems biologists to detect discrepancies between candidate mechanisms and reality

that might otherwise go unnoticed. However, in order to be powerful as heuristics, small models rely on, at times crude, idealizations. For this reason, experimentalists might reject such models as 'distortions' and consider the claims made by systems biologists as irrelevant to the plausibility of their proposals.

3.2.2. Large models

In areas where detailed knowledge about underlying causal structures has been accumulated, systems biologists sometimes attempt to build larger and more realistic models. A model of this kind is usually not restricted to one single phenomenon, but spans several interconnected processes that before were investigated separately. Examples are the comprehensive model of cell cycle regulation in budding yeast (Chen et al. 2004) and, more recently, the whole-cell model of Mycoplasma genitalium (Karr et al. 2011). The aim of such large scale modeling projects is often to check the sufficiency and consistency of existing explanatory schemes. Here is how the group behind the cell cycle regulation model describe the rationale behind their project:

> For complex, interconnected networks (...) it is impossible to anticipate all the consequences of multiple mutations by undisciplined, "hand-waving" explanations. To be certain of the sufficiency and consistency of the mechanism, we must create a well-defined mathematical representation of the molecular interactions and demonstrate that the model fits all (or most) of the relevant data. (Chen et al. 2004, p. 3858-3859)

This confirms the basic idea that computational approaches are used to reduce the influence of 'undisciplined' human reasoning. Aside from overcoming the limitations of qualitative and sequential reasoning, large models also have the potential to correct for biases that decomposition and localization introduce at a higher level. Differently from the small models discussed in the previous section, they take into account the interplay between parts that are usually studied as separate, quasi-independent modules, thus recomposing the system as a collection of integrated mechanisms. Studying such models, systems biologists sometimes discover unexpected behavior at the interface between conceptually separated mechanisms that can subsequently be investigated more thoroughly by experimental means. Karr et al., for example, found that

the length of the cell cycle in simulations of their whole-cell model is subject to an 'emergent' regulation that can be understood only when taking into account the build-up and depletion of the pool of nucleotides.

Large models are thus potentially very useful tools for aiding biologists in the development and correction of comprehensive accounts of integrated cellular systems. Nevertheless, they are far from being automatized and unbiased tools to check proposed mechanistic schemes. As in the case of small models, there is no one-to-one translation from wiring diagram to mathematical description. Moreover, the problem of having to specify many free parameters is usually aggravated as the size of the model increases. The amount and the right kind of empirical data that would be needed to accurately determine the free parameters is often not available, and systems biologists have to apply all kinds of tricks to keep the modeling problem well-constrained. Apart from the usual idealizations that were already discussed above, they are often forced to introduce strong assumptions about basic organizational features of the system under study. Karr et al., for instance, build their whole-cell model from smaller models that are assumed to be quasi-independent:

> Because biological systems are modular, cells can be modeled by the following: (1) dividing cells into functional processes; (2) independently modeling each process on a short timescale; and (3) integrating process sub-models at longer timescales. (Karr et al., 2012, p. 399)

However, they do not provide a justification for the assumption of modularity, nor for their particular choice of using a 1s timescale to model the processes independently. This is not to deny the role that large-scale models can play in testing mechanistic models, but in order to properly assess their value, it is important to be be realistic about their inherent limitations. The title of Karr et al.'s article ("A Whole-Cell Computational Model Predicts Phenotype from Genotype") suggests that the main goal of a large modeling project consists in successfully predicting or explaining cellular behavior based on the known underlying molecular structure. In the article itself, however, the authors emphasize the value of the model as a tool for discovery:

> [E]experimental analysis directed by model predictions identified previously undetected kinetic parameters and biological functions. We conclude that comprehensive whole-cell models can be used to facilitate biological discovery. (Karr et al. 20012, 389)

3.2.3. Network approaches

The small and large models discussed above are still committed to the decompositional strategy that also guides the traditional experimental approaches. They rely on the same general perspective of functional modularity, even though they introduce different strategies to tackle complexity within the modules they are studying. However, the system-wide study of the architecture of living organisms, enabled by the various 'omics' projects of data collection, suggests that underlying many of the behaviors of biological systems are large networks of interacting components that seem to go far beyond the simple schemes of sequential or modular organization with which human cognition feels comfortable. It is not obvious whether the strategies discussed so far can simply be scaled up to this level:

> [M]ost biological characteristics arise from complex interactions between the cell's numerous constituents, such as proteins, DNA, RNA and small molecules. Therefore, a key challenge for biology in the twenty-first century is to understand the structure and the dynamics of the complex intercellular web of interactions that contribute to the structure and function of a living cell. (Barabási and Oltvai, 2004, p. 101)

Many systems biologists believe that additional conceptual tools, in particular the tools of network theory, are necessary in order to understand biological systems at larger scales.

Network theory was developed in the 1930s, largely within the social sciences. It became more widespread in its applications when connections with mathematics, especially graph theory, were established in the 1950s. The basic idea is to represent a system, in a very abstract way, as a series of nodes that are connected by links standing for pairwise interactions or relationships. One of the aims of the theory is to find quantitative measures of network properties in order to classify different types of networks. Network theory gained considerable popularity across the scientific community after it was shown around the turn of the millennium that networks as different as the world wide web, electrical power grids, and metabolic networks share some unexpected features, such as the property of being scale-free (Jeong et al., 2000). If a network is completely random, most nodes have roughly the same number of links, or degree. In scale-free networks, by contrast, the degree distribution

follows a power law. This means that most nodes have only very few links, while there are a few nodes, called 'hubs,' that are highly connected. The fact that diverse types of networks share this non-random property led many scientists, and notably biologists, to expect that the general concepts of network theory had the potential to reveal deep underlying principles and might increase our understanding of large complex systems (Keller, 2005).

The study of universal properties of networks, however, did not turn out to be as fruitful as expected in biology. Nevertheless, many systems biologists hope that more specific network approaches, that elaborate on concepts from network theory, will lead to important progress in the study of complex biological systems:

> By itself, the fact that a network has scale-free properties is of limited use to biologists. Power laws occur very widely in nature and can have many different mechanistic origins. If we wish to obtain testable biological insights, we must probe further into the substructure of the network. (Bray 2003, 1865)

One way to gain insight about network substructure is the approach of network motifs, developed mainly by the group of Uri Alon at the Weizman Institute in Tel Aviv. He describes the aim of his work as follows:

> Our goal will be to define understandable patterns of connections that serve as building blocks of the network. Ideally, we would like to understand the dynamics of the entire network based on the dynamics of the individual building blocks. (Alon, 2007, p. 27)

On this view, networks are not unfathomable assemblages of interconnected nodes, but consist of substructures that are situated somewhere between the level of the single node and the level of the whole network. These building blocks reveal themselves through recurring patterns of connectivity, or 'motifs.' The key idea behind the search for motifs in a network might be called a 'reverse engineering' strategy. The goal is not to determine the structure of molecular interactions underlying a particular behavior as in the approaches discussed in the previous sections. Instead, one starts with the structure of the whole network and tries to make inferences on possible function. The search for motifs ideally begins with a complete description of the network's topology, that is, a map containing all the nodes and edges. Afterwards one applies a criterion of statistical significance to identify recurring patterns within

subgraphs of the network. In order to find such a criterion, the network under study is compared to a computer-generated ensemble of randomized networks. Network motifs are those patterns of connections that are found much more often in the real network than in the randomized networks. Their overrepresentation suggests that they are biologically meaningful and might play specific roles in the network. Alon's group has shown that there are a few types of motifs, such as the autoregulation motif or the feed-forward motif, that are highly significantly enriched in biological networks compared to random networks. Detailed mathematical analysis of individual motifs can suggest possible functions which are typically understood in analogy with signal-processing elements known from electrical engineering. A further step consists in investigating how the different motifs are integrated within the whole network. For the case of the sensory transcriptional network of the bacterium *E. coli*, Alon's group finds a relatively simple layer structure that suggests to think of the network as a kind of computer that consists of integrated signal-processing units: "Overall, the rather simple way in which the network motifs are integrated makes it possible to understand the dynamics of each motif separately, even when it is embedded within larger patterns" (Alon, 2007, p. 90).

Differently from the modeling strategies discussed before, the network motif approach does not apply the heuristic of decomposition and localization. Even though the aim is ultimately to explain the behavior of a complex system, this behavior does not directly guide the investigation. Instead, the first step is to represent the structure of the complete system topologically and to look for peculiarities in this structure. But just as in the case of decomposition and localization, also here the success of the strategy relies on several assumptions about the system. The various criticisms that have been put forward against the idea of network motifs raise the question of whether the simple structures found in network representations point to biologically relevant features or are instead artifacts of our representational tools and the need to bring overwhelmingly complex systems within the conceptual grasp of the human mind. In particular, the criticisms reveal that the technical assumptions underlying the network motif approach are inextricably linked with the way the underlying biological system is conceived and represented. As pointed out by Artzy-Randrup et al. (2004), for example, the choice of randomized networks that serve as a null-model for motif detection makes implicit assumptions about the process of network evolution and has a significant effect on the results.

Solé & Valverde (2006, p. 419) go further in this direction, calling network motifs "spandrels of cellular complexity," thereby alluding to the famous critique of the adaptationist program by Gould & Lewontin (1979). Just as evolutionary biologists sometimes get off target when trying to devise an adaptive explanation for every seemingly functional trait of an organism, systems biologists may be fooled by prematurely equating statistical with functional significance in their molecular data.

In summary, the approach of network motifs shows that there are alternative strategies to the study of complex systems that do not presuppose a functional decomposition. However, even though the application of an automatic statistical procedure seems at first glance as the hallmark of an unbiased approach, one has to take into account that the biological interpretation of the results depends on various assumptions that reveal its heuristic nature.

4. Conclusion

In this short overview I have tried to show some general characteristics of different computational approaches in systems biology. The epistemic problem of biologists can often be described as a search for mechanisms that explain phenomena of interest. Heuristics are strategies that facilitate these tasks of searching and explaining. However, they are not error-proof algorithms guaranteeing a correct solution to the problem. The 'human factor' enters in the form of certain assumptions about organizational features of system under study that are needed to break down the complexity of the epistemic task. Experimentalists are mostly concerned with the problem of search, and one might think that computational models are used to understand the dynamics of the mechanisms once their structural features have been identified. As we have seen, however, computational models are often used to engage in the problem of search as well. They are able to facilitate the discovery of mechanisms by overcoming some of the limitations of the traditional approach. However, they can only be efficient by introducing heuristic assumptions of their own. Modeling is always based on a number of pragmatic choices and idealizations that are made mainly in the interest of tractability. One might say that an advantage of computational approaches is that it forces modelers to make all of their assumptions explicit. However, as we have seen in the discussion of network motifs, the opposite risk also exists: biases are hidden because an

approach has the characteristics of an automatized procedure. Even though computational power substantially extends our cognitive abilities and allows us to overcome some of the biases of unaided cognition, the assessment of whether a particular method is appropriate in a given context and whether its assumptions are justified is left to us.

To conclude, the human factor remains a substantial and necessary ingredient of biological research. Generally speaking, the aim of scientists should not be to free their methods from bias, but to ensure that there are good mechanisms for error-detection and correction. Pursuing alternative strategies in parallel seems an efficient way of detecting bias, as long as scientists are ready to acknowledge the limitations of their own approaches and to appreciate the potential strengths of those of others.

REFERENCES

Alon, U. (2007). *Introduction to Systems Biology*. Boca Raton: Chapman & Hall/CRC.

Artzy-Randrup, Y., Fleishman, S.J., Ben-Tal, N., & Stone, L. (2004). Comment on "Network motifs: simple building blocks of complex networks" and "Superfamilies of evolved and designed networks". *Science* 305(5687), 1107.

Barabási, A.L., & Oltvai, Z.N. (2004). Network Biology: Understanding the Cell's Functional Organization. *Nature Reviews Genetics*, 5(2), 101–113.

Bechtel, W. (2011). Mechanism and Biological Explanation. *Philosophy of Science*, 78(4), 533– 557.

Bechtel, W., & Richardson, R.C. (1993). *Discovering Complexity*. Princeton: Princeton University Press.

Black, M. (1993). More about Metaphor. In A. Ortony (Ed.), *Metaphor and Thought* (2nd ed.). Cambridge: Cambridge University Press, 19–41.

Bray, D. (2003). Molecular Networks: The Top-Down View. *Science* 301(5641), 1864–1865.

Brown, K.S., Hill, C.C., Calero, G.A., Myers, C.R., Lee, K.H., Sethna, J.P., & Cerione, R.A. (2004). The statistical mechanics of complex signaling networks: nerve growth factor signaling. *Physical Biology* 1(3), 184–195.

Chen, K.C., Calzone, L., Csikasz-Nagy, A., Cross, F.R., Novak, B., & Tyson J.J. (2004). Integrative Analysis of Cell Cycle Control in Budding Yeast. *Molecular Biology of the Cell* 15(8), 3841–3862.

Downward, J. (2001). The Ins and Outs of Signalling. *Nature* 411(6839), 759-762.

Gould, S.J., & Lewontin, R.C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 205(1161), 581–598.

Gross, F. (2013). *The Sum of the Parts: Heuristic Strategies in Systems Biology*. PhD Dissertation, University of Milan, Italy.

Ideker, T., Galitski, T., & Hood, L. (2001). A New Approach to Decoding Life: Systems Biology. *Annual Review of Genomics and Human Genetics* 2, 343–372.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., & Barabási, A.L. (2000). The Large-scale Organization of Metabolic Networks. *Nature* 407(6804), 651–654.

Karr, J.R., Sanghvi, J.C., Macklin, D.N., Gutschow, M.V., Jacobs, J.M., Bolival, B., Assad-Garcia, N., Glass, J.I., & Covert, M.W. (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell* 150(2), 389–401.

Keller, E.F. (2002). *Making Sense of Life: Explaining Biological Development with Models Metaphors and Machines*. Cambridge: Harvard University Press.

MacLeod, M. (2015). Heuristic Approaches to Models and Modeling in Systems Biology. *Biology and Philosophy* (online first).

Morrison, M., & Morgan, M.S. (1999). Models as mediating instruments. In: M.S. Morgan, & M. Morrison (Eds.), *Models as Mediators*. Cambridge: Cambridge University Press, 10–37.

Resnik, D. (1997). Adaptationism: Hypothesis or Heuristic? *Biology & Philosophy* 12, 39–50.

Simon, H.A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society* 106(6), 467–482.

van Regenmortel, M.H.V. (2004). Reductionism and Complexity in Molecular Biology. *EMBO reports* 5, 1016–1020.

Wimsatt, W.C. (2007). *Re-engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge: Harvard University Press.