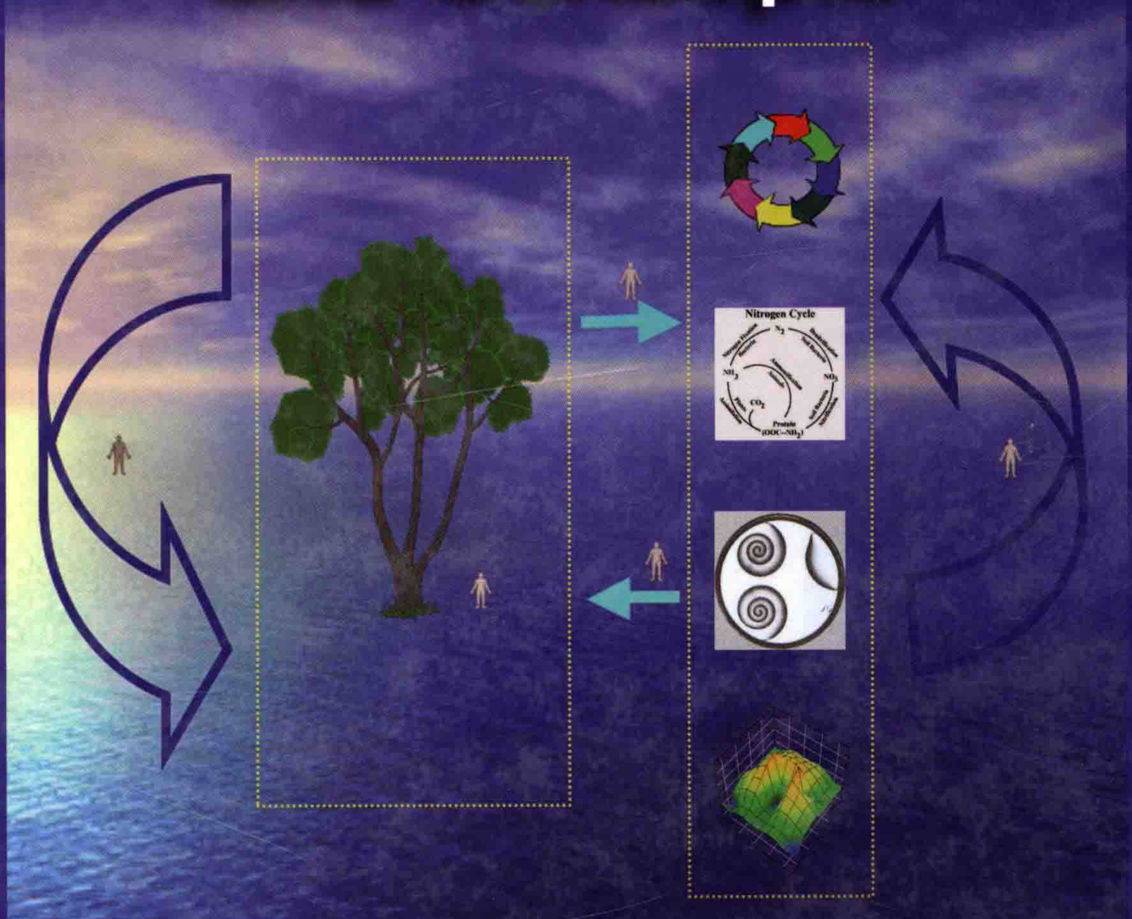
SYSTEMS BIOLOGY

Principles, Methods, and Concepts



Edited by

Andrzej K. Konopka

CRC Press
Taylor & Francis Group

SYSTEMS BIOLOGY

Principles, Methods, and Concepts

Edited by

Andrzej K. Konopka

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an informa business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2007 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-8247-2520-4 (Hardcover)
International Standard Book Number-13: 978-0-8247-2520-4 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

SYSTEMS BIOLOGY

Principles, Methods, and Concepts

PREFACE

This text presents biology as an autonomous science from the perspective of fundamental modeling techniques. It is designed as a desk reference for practitioners of diverse fields of life sciences, as well as for these intellectually mature individuals who would, themselves, like to practice the art of systems biology in the future.

Albeit systems biology exists for well over two millennia, it has enjoyed a spectacular rejuvenation in the recent years. The computer has indeed been a major tool for systems scientists, including systems biologists, since at least the 1960s. This is probably a reason why most conceptual foundations of today's systems biology appear to be expressed with the help of terminology borrowed from logic (Chapters 1, 7, and 8), linguistic (Chapters 1, 7, and 8), theory of knowledge (Chapters 1, 2, 4, 5, 7, and 8), computer science (Chapter 1, 2, and 8), general systems theory (Chapters 1, 6, and 7), and dynamical systems (Chapters 2, 3, 5, 6, and 9). Because of the diversity of flavors of possible applications, the general modeling methods are presented from several different perspectives such as for instance biochemistry (Chapter 2), thermodynamics (Chapters 3 and 9), engineering (Chapters 7 and 8), and ecology (Chapter 5).

Each chapter has been carefully reviewed and edited such that it will most likely provide the reader with a factually and methodologically rigorous state-of-the-art tutorial, survey, and review of modeling convoluted (complex) organic systems. Of course it would be naïve to guarantee that all mistakes or misstatements are eliminated by the editing (most probably are). If any errors are left, I will certainly feel responsible for them and therefore I would greatly appreciate it if the readers could point them out to me by writing an e-mail or a "snail" mail.

At this point, I would like to extend my most sincere thanks to Jean-Loup Risler (University of Evry) and Laurie F. Fleischman (BioLingua™ Research, Inc.) for multiple reading, reviewing, and help with editing the scientific content of the chapters of this handbook. The task of meritoriously revising the final drafts of this book would have simply been impossible without their selfless contribution. I would also like to thank our outstanding acquisition editor Anita Lekhwani who not only initiated this project but also relentlessly motivated it until its fruition, with her enthusiasm and spectacular resolve. Finally, I also owe thanks to my organization, BioLingua™ Research, Inc., and its board of directors for their enthusiastic support of my involvement in this project, as well as for releasing me from the fund-raising duties and teaching assignments during almost the whole of 2005.

Andrzej K. Konopka
Gaithersburg, MD

Contributors

B. Andresen
Niels Bohr Institute
Copenhagen, Denmark

S. W. Kercel
New England Institute
University of New England
Portland, Maine

W. Klonowski
Institute of Biocybernetics and Biomedical Engineering and GBAF Medical Research Center
Polish Academy of Sciences
Warsaw, Poland

A. K. Konopka
BioLingua™ Research, Inc.
Gaithersburg, Maryland

J. Nulton
Department of Mathematics and Statistics
San Diego State University
San Diego, California

K. L. Ross
Department of Philosophy
Los Angeles Valley College
Van Nuys, California

A. Salamon
Department of Mathematics and Statistics
San Diego State University
San Diego, California

P. Salamon
Department of Mathematics and Statistics
San Diego State University
San Diego, California

J. H. Schwacke
Department of Biostatistics, Bioinformatics, and Epidemiology
Medical University of South Carolina
Charleston, South Carolina

R. E. Ulanowicz
Chesapeake Biological Laboratory
University of Maryland Center for Environmental Science
Solomons, Maryland

E. O. Voit
The Wallace H. Coulter Department of Biomedical Engineering
Georgia Tech and Emory University
Atlanta, Georgia

Contents

functionally robust but otherwise complex whole. It also remains true of other basic methods and concepts of life sciences (such as, for instance, embryogenesis) described in Aristotle's work [1].*

I should mention here that because of *semantic* and *pragmatic* ambiguity of the popular credo of scientific holism [3,4], "system is more than a sum of its parts," one can distinguish several coexisting paradigms of systems biology that generate different shades of meanings of words "system," "more," "sum," and "parts." In contrast to physics-related research programs, there does not need to be a "winning paradigm" in biology. Several different schools of thoughts, methodologies, and belief systems can easily coexist in *life sciences* and it is reportedly a rare event that one well-formulated paradigm would prevail (i.e., "win") over the others [4].

Despite an impressive, time-resistant success of the foundations of biology as an autonomous science, today's academic methodologies of science clearly pertain to mechanics and arguably to a few other branches of physics while completely neglecting the cultures of chemistry [5] and, even more so, of biology [4–8]. As a consequence, it has been erroneously assumed by many that by setting appropriate habits of experimentation and theorizing (external controls), it must be always possible to eliminate the influence of the observer's mind (internal controls) on scientific observations in every imaginable field of science, not only mechanics. In reality, this assumption has been only partly effective even in physics proper and, without a doubt, research in biology and chemistry continues to rely not only on external but also on internal controls at the researcher's disposal. It is fair to say that the more convoluted (complex) a material system is the more important the properties (activities) of its observers (modelers) appear to be.

Within similar lines I wrote in another text a few years ago [9]:

> … As most of us know, heuristic reasoning can be conclusive. Yet biologists face the problem of not being able to specify all the rules applied to derive conclusions. Nor are they able to list all assumptions on which those rules ought to operate. It seems that these inabilities are a reflection of the complexity of biological systems themselves. Biological phenomena are often represented by models that are still too complex to be described in a communicable manner. Further and further modeling is required until our observations can be communicated in a linguistically comprehensive way. The cascade of models gives us the advantage of creating "communicable reality" but does not help us to judge the evidence pertinent to "real" (i.e., not necessarily communicable) reality. To the contrary, the more advanced a model in a cascade is, the further is its "distance" (in terms of number of modeling steps) from the modeled system ….

We could perhaps add that because of the extraordinary complexity of living things and because of biologists' dependence on natural language, biology has been considered a "descriptive" science for centuries. It has therefore not been unusual for biologists to believe that the actions of observers are as important for the outcomes and interpretations of findings as the actual observations are assumed to be. An additional reason for the scientific merit of taking into account an observer is the fact that our human means of observation (such as senses) do not need to be appropriate for even detecting, not to mention their modeling, the complex natural phenomena we would be

---

* Most of Aristotelian foundations of modern biology are described in his work entitled *Physics*. It is therefore relevant to notice that what Aristotle called "Physics" should not be confused with today's name of the science called physics. For one thing, Aristotle's attempts to study physical and astronomical phenomena have reportedly been quite lousy and have long been forgotten. In contrast to that, his work on foundations of biology continues to inspire generations after generations of biologists ever since his written testimonials have been known. In the absence of better terminology, Aristotle called "Physics" all kinds of studies of Nature; particularly living things such as plants and animals that today would belong to biology. Incidentally, it is believed that the term "biology" did not exist until Jean-Baptiste Pierre Antoine de Monet, Chevalier de Lamarck (known to most of us as Lamarck) mentioned it for the first time in his famous introduction to his "Philosophie Zoologique" published in 1809 [2]. Since this first published mention of the word *biology* refers to a title of an unfinished manuscript that Lamarck begun writing between 1800 and 1801, it is believed that he invented the term *biology* around 1800. (Some other accounts credit Lamarck and Treviranus for independently coining the term *biology* in 1802). Anyway, Aristotle could not possibly be aware of the work undertaken 2200+ years after his death and therefore he included today's biology in his "Physics."

potentially interested in observing if we were able to. In this situation we rely on models that (again) may be very distant from the actual natural objects, phenomena, and processes we should be observing. This in turn brings us to yet another good reason for taking into account an observer: the attributes of models we are capable of observing may or may not be adequate to the actual natural systems we should be observing.

It follows from the foregoing comments that, in an abstract characterization, systems biology can be seen as studies of properties of pairs, each consisting of an *observer* and a subject of observation (an *observed*) [10].

1.2 GENERAL SYSTEMS THEORY

General systems theory [11–13] (whose variants are also known under the names of *cybernetics, systems research, or systems thinking*) is an academic culture that promotes the *functional organization* as a primary concept instrumental for systems definition and description. In most general settings, it is based on the assumptions that:

1. Systems behavior (functioning) as a whole is governed by general rules (or laws).
2. The rules do not need to be expressed in terms of properties of the components (parts) of the system and ideally would not have much *relevance* to the laws followed by the parts.

Traditionally, the general systems theory has been known in four complementary practical incarnations (paradigms) [10] that explored different ways of overcoming misgivings of reductionist methodologies (see section 1.4 of this chapter), as well as approaching a working definition of complexity:

1. Chaos theory [14]
2. Cellular automata [15]
3. Catastrophe theory [16]
4. Hierarchical systems [17]

The first three of the foregoing incarnations of general systems theory (chaos, cellular automata, and catastrophes) are all variants of dynamic systems, each of which in turn can be represented by a finite set of differential equations. The fourth incarnation of systems biology paradigms, hierarchical systems, appears to constitute a focal point of methodology. On the one hand the postulate of existence of hierarchical systems and the concept of hierarchy constitute a condition of clear thinking (condition 3 in section 1.2.1) while, on the other hand, the hierarchical organization can be used to assume that functioning (action) at one level of hierarchy does not need to be reducible to functioning at another level.*

1.2.1 PRINCIPLES OF CLEAR THINKING

One aspect of the appearance of complexity is our complete dependence on the object language in which observer's causal questions (what? how? and why?) are formulated and the answers are communicated. The object language of scientific observations [4,5,10,12,18–21] is meant to be simple but the observed situations almost always appear to be complex, at least *a priori*. Therefore,

---

* Of course the opposite assumption, stating that properties at one level can be derived from behavior of things on another level, can be and sometimes is made in the spirit of causal reductionism (see discussion in section 1.4 of this chapter.) Our point here is that the framework of hierarchical systems allows one to choose between reductionist and nonreductionist principles of methodology. Whether the choice is really essential or not is debatable and the answer clearly depends on the specific nature of the problem.

the observer is forced to actively ask and answer questions that pertain to the following *principles of clear thinking* during the act of observing as well as before and after it:

1. Principle of *formal correctness* — effective reasoning in science should conform to the rules of logic (i.e., should be formally correct)
2. Principle of *causality* — every scientifically explainable object, process, or phenomenon (a fact) can be thought of as materially or logically entailed by a cause (also a fact, object, or process)
3. Principle of *hierarchical organization* (a postulate of the existence of hierarchy of things) — every scientifically explainable object, process, or fact belongs to a hierarchical organization (a hierarchical system) and, at the same time, contains things organized hierarchically.
4. Principle of *logicolinguistic fractionability* — every scientifically explainable system can either be represented as a subsystem (or a part) of a plausible (or well-defined) larger system or be partitioned into subsystems whose properties are believed to be plausible or well defined. To avoid the so-called *infinite regress,* one additionally assumes the existence of the minimal parts (logicolinguistic atoms) that cannot be broken into smaller parts.
5. Principle of *material adequacy* [21,22] — all concepts and linguistic constructs (particularly definitions) used to explain observations of natural systems must not only be true but must also correspond to factual reality to a maximum degree possible. Our explanations must be as relevant and as pertinent as possible to the actual observed process or phenomenon that we try to explain. (Trivial or irrelevant truths should not be acceptable as valid explanations).

The foregoing principles are arguably the rules for the observer's "mind only" because they cannot be effectively applied without semantic and pragmatic interpretation.* That is to say even if a hard-core reductionist believer in materialism would like to see the parts and hierarchies right there in nature (as some extreme materialists do), it would not really matter because such constructs are only possible within a meta-language of reasoning performed by the observer's structuring mind. Ergo, the need for semantic and pragmatic interpretation of the rules of clear thinking are one of the reasons for systems biology to be concerned with the scientifically viable methodology of modeling, as well as with the distinction between a natural system and its model.

1.3 THE CONCEPT OF TRUTH IN NONDEDUCTIVE SCIENCE

About 2 years ago, I wrote [22]:

> In everyday practice of science we assume that when an individual communicates results of an experiment he or she intends to tell the truth. We need also to assume that all scientists intend to be truthful to the best of their human capacity and therefore their scientific communications should not contain intentional lies. Nor should their messages be intended to misguide or otherwise mislead others into believing that something is true that in fact is not so. In other words we assume that individuals whose craft is science choose to be truthful and honest and do in fact pursue the virtue of truthfulness to the maximum of their capacity in every act of scientific communication.

Even with these assumptions in mind, the business of approximating truth is intricate and elusive. First of all, we need to follow pragmatic principles of clear thinking (see section 1.2.1),

---

* The principle of formal correctness (principle #1) is perhaps an exception here. There exist formal systems (such as, for instance, propositional calculus in logic) whose syntax is sufficient to account for their own semantics.

which often affect the nature of questions we are able to ask. That is to say, it cannot be taken for granted that the questions we are capable of asking are most adequate regarding our interest in the answer. Second, we are forced to use a subset of natural language to formulate our queries and to communicate with each other. Ergo, we often need a linguistic convention that would define a meta-language to specify the (object) language designed for communication.

1.3.1 GRAND THEORIES OF TRUTH

Historically, we know of at least three different grand theories of truth:

1. Correspondence theory: What we say or see is true if and only if it corresponds to the factual reality. Every pursuit of this doctrine requires, from the outset, that we take into account the possible discrepancies between real facts and their models.
2. Coherence theory: What we say or see is true if and only if it conforms to an existing system of knowledge without contradictions. The validity of this theory depends on our capacity to make sure that we have taken into account all the existing knowledge (an obvious overstatement of our potential).
3. Pragmatic (utility-based) theory: What we say or see is true if and only if it is being adequately used. One technical problem with this doctrine is due to our ability (or lack thereof) to determine adequacy of usage. Another problem is integrity (or lack thereof) of the users. That is to say, concepts and assertions can be used with or without regard to merits of such utilization as long as it (the utilization) brings social, reproductive, political, or any other merit-unrelated success to the users.

There also exist theories of truth that are blends of the foregoing three "pure" doctrines. For instance, we can often encounter in science the following tacit doctrines:

4. Truth by convention (naive epistemological idealism): We agree to believe that certain general assertions are true and do not require the evidence to demonstrate this fact. (One classical example is an axiomatic system in mathematics, such as, for instance, Euclidian geometry, where the axioms are assumed to be true theorems.)
5. Truth by nature of things (naive ontological realism): Whatever we believe about the observed relation between objects or phenomena or processes, the relation is true or false in objective reality and has nothing to do with the observer's ability (or lack thereof) to determine that.

In formal logic and most of mathematics (i.e., in deductive sciences), we employ a formal language (such as, for instance, predicate calculus or arithmetic) to determine if a given specific statement formulated in this language is true or false. All we need to do is to check correctness of use of the rules of syntax of the language. In this respect, the theories of coherence (theory 2 above) and convention (theory 4 in our list) are favorite doctrines of truth in deductive science. The coherence theory and conventionalism are also doctrines of choice in various disciplines of computer applications that require handling of large data sets (such as for instance database searches, data mining, and pattern acquisition). The predominant paradigm in these fields is the creation of standards for judging the quality of data or software. (Presumably, any standard is better than no standard at all.)

The schemes of arriving at possible truth in nondeductive science appear to differ from field to field. For instance, most of physics and astronomy seem to rely on simple induction — derivation of general rules of system's activity from instances of specific processes. On the other hand, chemistry and biology — in a very general view — seem to require more than a simple induction to validate scientific claims. This not-so-simple induction along with other complicated modes of inference is often called *pragmatic inference* [5,9] and will be discussed briefly in the next section. It is not clear which of the five theories of truth will be used by individual researchers. It looks as though truth by correspondence (theory 1) is a favorite doctrine in systems biology and other fields of science* because of the built-in distrust in our models. In fact, the correspondence theory requires that we verify the fidelity, accuracy, and/or precision with which the models represent nature (i.e., the factual reality).

1.3.2 DEDUCTION, INDUCTION, AND PRAGMATIC INFERENCE

Forming conclusions from premises is called *inference*. One can distinguish several types (modes) of systematic inference. Their nonexhaustive list includes:

1. Deduction (from general to specific)
2. Induction (from specific to general)
3. Pragmatic inference
   a. Abduction (finding the most plausible explanation available)
   b. Consilience (transcending different sequences of inferences)

Formal logic, as we know it today, can only deal with deductive inferences and to some extent with induction (such as inductive proofs of properties of integers). All other types of systematic inference that appear similar to an intricate (i.e., not simple) induction are often referred to collectively as pragmatic inference [9,23,24]. They cannot be easily formalized and thereby require language and domain-of-application–dependent conventions that would define their semantics (syntax alone will not suffice).**

Inference via deduction always leads to specific conclusions from general premises. For instance, the implication: "every man is mortal, I am a man, and therefore I am mortal" is an example of deduction. The general scheme of deduction is:

1. Every object A has a property B.
2. Therefore, a given specific individual object A has a property B.

Induction is a derivation of a general conclusion from specific premises. The scheme of inductive inference is:

1. Every individual object A observed thus far has a property B
2. Therefore every A has property B

The question of what the phrase "thus far" should exactly mean has been hotly debated for centuries and is known under the name the *Problem of Induction* [25,26]. If we see the night following the day, say 1000 times, does it mean that the night *always* follows the day? Or perhaps, in order to decide that it is always a case, we need to see that happening 10,000 times instead of 1000? What about 10,000,000 observations? Will we be more certain of the outcome than we would have been with only a 1000 events?

---

* Mechanics and some other mathematics-dependent fields of physics (such as quantum theory) may be the exceptions here. In these fields, a naïve belief in reality of models (theory 5) seems to prevail over correspondence to factual reality. This is one of the nongeneric situations addressed by Hertz–Rosen modeling relation (see section 4 of this chapter) in which the surrogate formal system is identified with the modeled natural system. In typical situations regarding modeling nature, the two systems (natural and surrogate) should be kept separate.

** That said, the vocabulary and syntax of predicate calculus and quantification can be used as stenographic shortcuts in meta-languages describing all kinds of inferential schemes, including those of pragmatic inference.

Ever since Kant [27], the general answer to these kind of questions is to use reason and experience together in determining the number (and kind) of observations and then transcend all we know plus the outcomes of these observations into a new tidbit of knowledge. That has been one aspect of Kant's and his followers' transcendental methodology. It has also been an underlying principle of heuristics involved in planning experiments in all fields of science since, at least, the Hellenic period in ancient Greece.

The question of how to effectively do the transcending of our existing knowledge in the light of new observations and new inferences is a subject of specific methodology of each field of science and sometimes a style of individual researcher. However, there exist a few general guidelines that can be summarized under the name of pragmatic inference (blend of few types of inference plus a few pragmatic rules of transcending knowledge). Two known examples of types of pragmatic inference, abduction and consilience, are described below.

Abduction, also known as inference to the best explanation, is a derivation of the most plausible premises (explanations) from known alleged consequences. The most general scheme of abduction is:

1. B is a set of observations
2. A, if true, has a good potential to be a valid explanation of B
3. No hypothesis other than A appears to explain B as well as A does.
4. Therefore, A is probably true

C.S. Peirce who coined this term [28] has noticed that abductive reasoning can and often does violate principles of formal logic and set theory. An example of typical abductive reasoning is:

Sentence 1: My bank has many $20 bills.
Sentence 2: I have a $20 bill in my pocket.

Therefore

Explanation: I obtained my $20 bill from my bank.

The obvious logical problem with this explanation is the fact that my bank does not need to be a unique source of $20 bills. We could make the explanation stronger with the help of additional evidence (such as, for instance, nonexistence of other than my banks in my geographic area or some peculiar additional characteristics of the specific $20 banknotes) but generally abductive reasoning tends to violate rules of logic.

Despite the logical weakness of abduction, it is a popular tool of providing plausible explanations. Perhaps the most dramatic examples of acceptance of logically flawed abductive conclusions are superficial statistical inferences such as, for example, this:

Statement 1: A large set (population) X contains a large number of red objects.
Statement 2: A not so large (smaller than X) set Y contains a significantly large number of red objects.

Therefore

Conclusion: Y is a subset (sample) of X (population) or, in other words, X contains Y.

It is obvious that without factual, empirical evidence for Y being a subset of X, the conclusion is false and a statistics done on Y does not need to correspond to statistical properties of X at all. In particular, the elements of Y do not need to constitute a sample drawn from the population X. Despite the misgivings of the foregoing erroneous sampling, the flawed "statistical" inference is commonly performed to demonstrate statistical support for specific experimental conclusions. In other words, abduction is and will be with us for the time to come but we need to follow Peirce's advice and exercise methodological prudence while using it, at least by making sure that sampling for our own statistical studies is done correctly.

Consilience, as we understand it today, is a synergistic combination of two or more (originally unrelated) sequences of inferences that can be transcended into a useful new concept or hypothesis. Statements or intermediary conclusions from one sequence can synergistically enlighten another (otherwise unrelated) chain of inferences and even (abductively) affect the conclusions.* The important aspect of consilience is a semantic and/or pragmatic interpretation (that may or may not be biased by further transcending by the observer's "structuring mind") of observations and syntax-based inferences. In many respects, consilience can be seen as abduction and, even more so, as a way of *knowledge integration* via creation of novel (or just useful) abstract entities. From this perspective, one can distinguish the following types of consilience:

1. Data integration via *naming* (or renaming) and adjusting *reference* — Objects, phenomena, and processes are bound together with concepts as well as relations between concepts. The result of such blend is then phrased in a subset of natural language, such as a jargon of a given discipline. An appropriately implemented metaphor, for instance, can facilitate understanding of a group of phenomena in light of an existing or a brand new conceptual framework alike. Similarly, a properly chosen analogy can catalyze understanding by bridging the phenomena that were unrelated to each other before.
2. Data *compression* via *generalization* — A set, usually a very large one, of individual factual observations can often be described in a brief manner with the use of knowledge from outside this data collection. The best examples here are trajectories of planets that, once discovered, compress the detailed knowledge of a huge number of otherwise unrelated positions of a celestial body in space. Similarly, chemical equations bring our understanding of chemical reactions to an exponentially more satisfactory level than a huge set of individual descriptions of experiments could afford. Numerous other examples, such as gas laws, include the use of compact mathematical equations to describe (model) factual situations that without such formalism would merit a lengthy and often inconclusive description.
3. Action (activity) *iconization* — Advanced use of symbols can, and often does, enhance or redirect understanding. All kinds of schemata (including the overwhelming presence of hierarchies in our definitions and reasoning) can symbolically represent a large body of our knowledge (new and old alike) without any linguistic description of what exactly constitutes this enlightened wisdom. Even a single symbol charged with a distillation of a huge body of shared cultural experience, such as a trash can on desktops of most personal computers, can be used to enhance or redirect understanding without any exhaustive explanation. The essence of iconization is abbreviating a conceptual mindset shared by all potential observers with common experience. The symbols used as icons must be interpreted in the same way by all their users or else the icons will not have

---

* The original concept of consilience has been coined by the "gentleman of science" William Whewell (1794–1866) in his 1840 treaty on the methodology of inductive sciences [29]. In this original version, he discussed only consilience of inductions as a source of unexpected discoveries that could not be attained by simple induction alone. However, the pragmatic principle of synergistic (or inhibitory) influences of sequences of inference need not be limited to inductive inference and for all practical purposes such limits do not exist. As a matter of fact, Whewell's understanding of the notion of induction was reportedly [30] much larger than the standard meaning of this concept would imply. Whewell's "induction" should be understood, in today's terms, as a knowledge integration with the help of old and new facts and concepts configured in an effectively knowledge-enriching manner. Consilience, then, is a lucky combination of such "inductions," which brings a quantum leap of understanding a phenomenon or a process or a class of processes.

their knowledge-summarizing power (and thereby will not be icons in other than an open-interpretation, artistic sense).*

1.4 REDUCTIONISM VS. HOLISM: A REAL ISSUE OR A MERE DETRACTION?

Some time ago I characterized the naive version of reductionism in today's "postmolecular" and "postgenome" biology as follows [10]:

> According to the common-sense, naïve, interpretation of reductionism life should be fully explained via exhaustive studies of systems of differential equations (dynamic systems) but very few, if any at all, biologists would see the point in following this research programme. An extreme version of this paradigm adopts an assumption that phenotype can ultimately be modeled by dynamic systems (i.e. systems of differential equations in which rates, forces, are time derivatives). In a little less extreme version of naive reductionism genes are assumed to prescribe (in an unknown way) the epigenetic rules (also unknown) that in turn control interactions (of unknown nature and number) of proteins (of unknown kind and function) with each other as well as with other (also unknown) ligands.

As we can see from this quite accurate description, naive reductionist agenda, which has been so successful in physics of the past three centuries, is not very convincing in biology in general (not even in the physics-friendly "postgenome" biology).

That said, I hasten to mention that the long-standing debate between reductionism and holism has been complicated by the intensity of political agendas of the most recent two millennia (including the 20th century). Not only are the general methodological programs of natural science difficult to express in language but also the unforgiving political pressures in different historical periods prevented their scholars and savants from having any opinion at all. The lucky ones who actually could express their opinions about their methodological worldviews often did not survive this experience or, if they have, they tended to revoke their original claims under the stimulating ambience of prison cells, torture chambers, or fear thereof. The last few decades of the 20th century went a bit easier on opinionated scientists talking of the glory of reductionism. However, it would be a mistake to say that rare instances of unemployment among (Western) life science faculties as well biologists' "vacations" in remote parts of former Soviet Siberia had nothing to do with attempts to encourage silence about the holistic (systemic) methodological attitudes. In fact, silencing (or enhancing) the scientists' general worldviews continues to go on all the time for the reasons that have nothing to do with science or even with the merits of the silenced (or enhanced) opinions themselves. In the face of all this methodological and ideological mess, a good question to ask is: What is the actual problem to be addressed with the help of learned opinions about the worldview?

As far as systems biology is concerned, the pivotal issue is reportedly a possibility of novelty of systems behavior regarding its ambiance (generalized material and symbolic environment). In principle, the occurrence of new, unexpected systemic characteristics (novelty) could be induced via either emergence or anticipation or both. Emergence is an apparently spontaneous generation of novel systemic properties that could not be predicted from the properties of components (parts) of a complex system alone. Anticipation is characteristic of the so-called anticipatory systems [31]** that contain predictive models of themselves as well as their ambiance and thereby can

---

* Specialists in a given field experience similar observations and use similar tools of interpretation. Because of that, they can effortlessly communicate their opinions to each other but are often misunderstood by nonspecialists.

** An anticipatory system is "... a system containing a predictive model of itself and/or its environment, which allows it to change state at an instant in accord with the model's prediction pertaining to a later instant." [31]. For the outside observer, the net effect of emergence should be the same as that of an anticipatory behavior (i.e., the apparent novelty) but the actual explanation of the phenomenon of novelty is entirely different.

unexpectedly, for the observer, change their current activity (behavior if you will) in response to the predicted future states.

The intense and often politically charged debate mentioned earlier in this chapter was to a great extent due to a fight between believers and nonbelievers in genuine novelty induced by emergence [32]. But were the reasons for this fight or debate scientifically meritorious? Anybody who has ever seen plants emerging from the seeds planted long ago does not have doubt that emergence is a fact. Anybody who ever made a chemical synthesis or even saw a chemical reaction occurring in a test tube does not have doubt that emergence is a fact. As a matter of fact, the phenomenon of emergence appears to be so common in nature that it should pass even the most severe test of systematic doubt. Why then is disbelief in spontaneous generation of novelty (such as emergence) possible? Well, because admitting the fact of emergence (and/or anticipation for that matter) sometimes contradicts other beliefs that their practitioners may find important for right as well as wrong reasons. One of the best characterizations of this situation is given by the dean of modern biology, Mayr [32], while describing properties of emergence:

> During its long history, the term "emergence" was adopted by authors with widely diverging philosophical views. It was particularly popular among vitalists, but for them, as is evident from writings of Bergson and others, it was a metaphysical principle. This last interpretation was shared by most of their opponents. J. B. S. Haldane (1932:113) remarked that "the doctrine of emergence ... is radically opposed to the spirit of science." The reason for this opposition to emergence is that emergence is characterized by three properties that appear at first sight to be in conflict with a straightforward mechanistic explanation: first that a genuine novelty is produced — that is, some feature or process that was previously nonexistent; second, that the characteristics of this novelty are qualitatively, not just quantitatively, unlike anything that existed before; third, that it was unpredictable before its emergence, not only in practice, but in principle, even on the basis of an ideal, complete knowledge of the state of the cosmos.

Reductionism is a collective label intended to mark several worldviews and general research strategies (sometimes distinctly different from each other). Selected aspects of these strategies sometimes indeed reside behind the actual day-to-day methodology of scientific research but sometimes are erroneously believed to motivate scientific methods.

One can roughly distinguish at least the following four variants of reductionism [10,33]:

R1. *Ontological Reductionism* — a worldview that the whole is strictly "nothing but" the sum of parts. One can clearly distinguish two extreme traditions here:
  R1a) A doctrine of *reductive materialism*: All properties of the system should be completely derivable from properties of its parts and therefore emergence or anticipation, as a generation of surprising novelty, could not happen at all.
  R1b) A doctrine of *reductive idealism* sometimes referred to as *physicalism*: All parts of the system fit perfectly the grand "plan" or "design" according to which nature is organized. There is no room for novelty as an attribute of nature. Any impression of novelty should be due to our ignorance about some behaviors of the parts but the systems characteristics should still be (ultimately) a combination of properties of the smallest components.* Ergo, if we happen not to satisfactorily explain the system from the properties of its alleged smallest parts, we would need to look for errors in identification of components or properties and then repeat the process of explaining over and over again. Not surprisingly, reductive idealism is almost identical to the strong version of holism (see the sequel to this section).

---

* Physicalism is *de facto* a mixture of ontological and causal reductionism because it is based on the assumption of bottom-up, upward, causation and the existence of the smallest components (atoms) that cannot be further divided into parts.

Ontological reductionism is probably extinct from science by now but it had played a constructive role as a belief system that fueled the development of mechanics, mechanical devices, and even medicine (anatomy, morphology, cytology) for several centuries.

R2. *Methodological Reductionism* — a research strategy that explores a concept of substituting a given system with a surrogate system. In this variant of reductionism, there should be no problem with accepting emergence as long as it could be explained by the properties of a surrogate system. The methodological version is pragmatically useful in science because it leads to a clear-cut distinction between a natural system (ontological entity) and its description (epistemic entity). This distinction is in fact reflected in a meta-language describing the *Hertz–Rosen modeling relation* [8,31,34] discussed in several places in this volume (including this chapter). In the case of convoluted (complex) systems, the boundary between modeled system and its model (the epistemic cut [35]) is not easy to find and, when found, often leads to paradoxes or inadequate theories. That is one of the reasons why a singe (largest) model cannot possibly be sufficient for an adequate description of convoluted (complex) systems. To approach a satisfactory systems description, a multitude (but an unknown number) of complementary models is needed.*

R3. *Epistemological Reductionism* — a view that the whole can (in principle) be a sum of parts subject to adequacy of definition of "parts" and "sum." Again, there should be, in principle, no problem with emergence as long as it could be derived from properties of an adequately selected set of parts. This version of reductionism is often associated with methodology of physics expressed, for instance, in modeling biological systems in terms of dynamical systems.

R4. *Causal Reductionism* — a view that upward causation ("the whole is there because of the parts") should be legitimate but downward causation ("the parts are there because of the whole") should be forbidden. Here, a genuine emergence is impossible unless we give up the property of an absolute novelty.

Causal reductionism can be seen as a special case of the epistemological one. It is clear by now that this version of reductionism fails in a most visible way in biology after being extraordinarily successful in physics since Newton. An obvious reason for this failure in the case of complex biological systems is the fact the concept of function, and thereby an idea of a goal of the system's activity, is in fact an important ingredient of a plausible biological explanation. A programmatic rejection of the entire idea of downward causation must upset any methodology, even a reductionistic one, which aims at an explanation of system's behavior in functional and not only structural terms. Living things appear complex because they are in fact convoluted in terms of their function.** An orthodox adherence to causal reductionism could arguably be methodologically sound for studies of structure at best. However all attempts known to this author (including his own attempts) of explaining biological function in terms of structure alone lead to inadequate or even irrelevant (trivial) explanations of their behavior (functioning).

---

* The nonexistence of the largest model and the need for a multitude of complementary models is the actual definition of a complex system given by Robert Rosen [8,34].

** That is why the concept of functional *complexity* in biology needs to be separated from the notion of mere complication due to a very large number of objects and relations between them. We often refer to this latter, combinatorics-related, concept as *algorithmic complexity* [36–38] or *compositional complexity* [9,23,39,40] but in the context of systems biology the word "complication" would be much more accurate. On the other hand, the concept of *convolution* reflects well the functional, causal, or inferential intricacy we are facing while modeling biological systems or interpreting their models. In other words, the property of being convoluted (and not the property of being complicated) is an essence of complex biological systems.