"Less Is More" and the Art of Modeling Complex Phenomena

Simplification May But Need Not Be the Key to Handle Large Networks

PETER SCHUSTER

Peter Schuster,
Editor in Chief of Complexity,
is at the Institut für Theoretische
Chemie der Universität Wien,
A-1090 Wien, Austria;
E-mail: pks@tbi.univie.ac.at

In the October 21, 2005 issue of the magazine *Science* a perspectives article by Stefan Bornholdt addresses the problem of the proper level of details in the description of complex systems [1]. "Less is more" is used in this article to encourage the usage of highly simplified dynamical elements for modeling large and complex nonlinear systems. The special case considered is modeling of large genetic networks, but the problem is much more general, arises often in science and beyond and, perhaps, deserves broader attention. Some older examples of "less is more" that have already reached a certain degree of maturity and common acceptance may be useful with respect to the recent revival of this paradigm. I shall present here two different problems from the interface of physics and chemistry, which were heavily debated in the past and for which consensus has been achieved by now, before returning to the burning biological questions.

The first example starts with Paul Dirac's famous comment [2], "… The underlying physical laws necessary for the mathematical theory of a large part of physics and the *whole of chemistry* are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. …" The reactions of the scientific public were extremely ambiguous: Quantum chemists used Dirac's statement as the figure-head for their many decades long search for better and better approximations to the Schrödinger equations, whereas the majority of experimental chemists were truly upset. The reason for the uneasiness also shared by other nonphysicists was certainly not only the overstatement of a then 27-year-old and somewhat arrogant physicist but also an intuitive feeling that quantum mechanics provides the tool to reduce and integrate chemistry into physics, thereby sacrificing chemistry's autonomy [3, 4]. Putting aside the philosophical questions, we are left with a pragmatic problem: Is quantum mechanics appropriate to describe molecules and chemical bonds for the chemist at the workbench? There are two reasons, among others, that suggest the application of a different, preferentially less sophisticated level of description: (i)

Some older examples of "less is more" that have already reached a certain degree of maturity and common acceptance may be useful with respect to the recent revival of this paradigm.

© 2005 Wiley Periodicals, Inc., Vol. 11, No. 2
DOI 10.1002/cplx.20107

Despite the spectacular progress of computational quantum mechanics that allows for incredibly accurate computation of structures and properties of small molecules [5, 6], calculations of large molecules are still far away from being satisfactory and, even more important, the predictive power of the full-blown quantum chemical approaches is rather week. In other words, we can calculate but we don't understand unless we crank up the highly sophisticated and costly computational machinery derived from the Schrödinger equation. (ii) Empirical chemical knowledge unlike the usage of the stationary Schrödinger equation unconsciously involves a time span of observation, and this matters when we discuss what we mean by chemical compounds. To give a naïve but illustrative example, dimethyl ether and ethanol, CH_3OCH_3 and CH_3CH_2OH, respectively, are two distinct chemical compounds, although they have the same Hamiltonian and are described by the same Schrödinger equation. Clearly, computing the energy landscape in the Born-Oppenheimer approximation and estimating the lifetime of the two isomers of C_2H_6O will undoubtedly reveal that the time scale for an interconversion of the two molecules is extremely long, and therefore we are not in danger that the ether is converted into more stable ethanol during an experiment. There are other pairs of isomers, for example, in nonclassical carbocations [7] that change structures too fast to be observed. Qualitative molecular theory, being a largely simplified and coarse-grained distillation of quantum mechanics for chemists, is often useful and makes successful predictions, although it is lacking the solid anchor in physics. The hybridization concept, for example, allows for many correct predictions of rough molecular geometries and the Walsh rules are likewise successful. The Woodward-Hoffman rules are a valuable tool for predicting the reactivity in certain classes of reactions. There is, however, one important fact to keep in mind: The qualitative picture fails inevitably when quantum mechanics is truly indispensable as it is, for example, in spectroscopy, in photochemistry, and in interactions of electromagnetic radiation with matter, in general. In addition there is no way around large-scale computation, if we are heading for predictions of quantitatively reliable and sufficiently precise results.

Qualitative molecular theory, being a largely simplified and coarse-grained distillation of quantum mechanics for chemists, is often useful and makes successful predictions, although it is lacking the solid anchor in physics.

My second case is another example from the crossroads of chemistry and physics: chemical reaction dynamics [8, 9]. Chemical reactions are commonly described at three main levels of sophistication [9]: (i) the qualitative or substance level, "What produces what under which conditions?"; (ii) the elementary step level, dealing with rate constants and their dependence on parameters like temperature, pressure, and other external conditions; and (iii) the full-blown chemical dynamics level where the interconversion of molecules are resolved to individual reactive collisions. Level (iii) provides marvelous insights into unexpected details of reactive quantum scattering and creates the link to computational quantum chemistry discussed in the previous paragraph. To mention just one illustration of such details: Simple reactive collisions, like $F + H_2 \rightarrow FH + H$, involve several atomic and molecular states and may be calculated now by quantum scattering techniques on multiple energy surfaces [8] and then, the computed results agree with molecular beam experiments. Paul Crutzen, who did the epoch-making studies on atmospheric chemistry and, in particular, ozone destruction by manmade pollutants [10], would have been completely lost if he had attempted to reach his goal on the quantum scattering level. All his success was based on consequent and precise level (ii) kinetic studies on vapor phase reactions. Sometimes even cruder descriptions between the elementary step resolution and the qualitative description are important. An illustrative example is the beautiful work on nonlinear chemical reactions in solution [11]. The famous oscillatory Belousov-Zhabotinskii reaction comprises some 20 or more elementary steps. In the "Oregonator" model developed by Richard Field and Richard Noyes [12, 13] these steps are cast into five overall reactions that allow for a perfect and accurate prediction of the course of the reaction and even of very subtle reaction details. Many examples could be added, which all demonstrate the more or less self-evident but nevertheless often forgotten fact: The proper model description of a complex system depends on both the context of the problem and the question one wants to ask.

Coming back to the initial problem concerning the proper level of description for complex biological networks we recognize a situation that is not very different from the two examples mentioned above. There are, for example, several levels of description for neural networks, I shall mention here only two of them: (i) The single neuron level, which is described in great detail by the famous Hodgkin-Huxley equation relating action potential and electric current in the neuron [14], and (ii) the highly coarse-grained level of neural networks that initiated a whole new area of computation (see, for example, the Hopfield networks [15]). At present both levels are still highly relevant: Level (i), because progress in the molecular biology of the neuron allows for a precise characterization of the molecular players in the Hodgkin-Huxley equation and calls for extensions of the original version to more realistic gating models, and level (ii), because we are still lacking a comprehensive theory for the emergence of collective properties in neural networks, in particular in the brain. With the current computational facilities it is also thinkable to combine both levels and to compute relatively large ensembles of Hodgkin-Huxley neurons. Here as well as above we have to face the "too much detailed" problem at the molecular level. On the other hand, when we are focusing on the role of individual classes of ion gates and specific neurotransmitters, the molecular level is indispensable.

Genetic and metabolic networks—*genabolic* networks might be a good name for the combination of both—are no exceptions of the rule [16, 17]. There are features for which the description by means of Boolean functions, as advertised in [1], is the most appropriate level to learn generic properties of signaling and regulation. When it comes to other questions, for example, the control of cellular activities by second messengers and hormones, the molecular level will be essential. The entire disciplines of computational systems biology and cell biology are in an exciting and very fast development. Despite impressive progress in the past few years several techniques have yet to be established and large-scale computations on the dynamics of whole cells and organisms will be impossible without specific advances in algorithms and their implementations. I see parallels to the development in computational chemistry, where the scientific questions were indeed completely forgotten for a few decades and people focused almost exclusively on the solution of computational problems. During such periods of technical progress, a reminder like Stefan Bornholdt's perspective that suggests not to forget the ultimate goals and to think about simpler approaches is undoubtedly in place.

The proper model description of a complex system depends on both the context of the problem and the question one wants to ask.

Is less more? The answer to the question, as I wanted to point out here, is subtle. It is "could be" rather than "yes," and whether or not it is true depends on the context and the problem to be investigated. The figure in the article in *Science* [1] distinguishes nicely four levels of description—single gene, small genetic circuits, medium-size, and large-scale genetic networks—and I think each one is justified in its own right. The art of modeling is to choose the proper degree of detail. One take-home lesson from the development of computational quantum chemistry, however, is that decades of methodological development, where everyone in the field focuses on the problem to compute faster and faster, larger and larger systems, may pay at the end, after the technical problems had been solved and the scientific questions come back into the focus of interest.

REFERENCES

1. Bornholdt, S. Less is more in modeling large genetic networks. Science 2005, 310, 449–450.
2. Dirac, P.A.M. Quantum mechanics of many-electron systems. Proc Roy Soc Lond 1929. A123, 714–733.
3. Benfey, T. Reflections in the philosophy of chemistry and a rallying call for our discipline. Found Chem 2000, 2, 195–205.
4. Lombardi, O.; Labarca, M. The ontological autonomy of the chemical world. Found Chem 2005, 7 125–148.
5. Pople, J.A. Quantum chemical models (Nobel lecture). Angew Chem Int Edit 1999, 38, 1894–1902.
6. Kohn, W. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. Rev Mod Phys 1999, 71, 1253–1266.
7. von Ragué Schleyer, P.; Maerker C. Exact structures of carbocations established by combined computational and experimental methods. Pure Appl Chem 1995, 67, 755–760.
8. Althorpe, S.C.; Clary, D.C. Quantum scattering calculations on chemical reactions. Annu Rev Phys Chem 2003, 54, 493–529.
9. Herschbach, D.R. Molecular dynamics of elementary chemical reactions (Nobel lecture). Angew Chem Int Edit 1987, 26, 1221–1243.
10. Crutzen, P.J. My life with O_3, NO_x and other YZO_x compounds (Nobel lecture). Angew Chem Int Edit 1996, 35, 1758–1777.
11. Sagués, F.; Epstein, I.R. Nonlinear chemical dynamics. Dalton Trans 2003, 1201–1217.
12. Field, R.; Körös, E.; Noyes, R.M. Oscillations in chemical systems. II. Thorough analysis of temporal oscillations in the bromate-cerium-malonic acid system. J Am Chem Soc 1972, 94, 8649–8664.
13. Field, R.; Noyes, R.M. Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction. J Chem Phys 1972, 60, 1877–1884.
14. Hodgkin, A.L.; Huxley, A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol 1952, 117, 500–544.
15. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA 1982, 79, 2554–2558.
16. Ptashne, M.; Gann, A. Genes & signals. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2002.
17. Fell, D. Understanding the control of metabolism. Portland Press Ltd.: London, 1997.