Part XIII

COMPUTER SIMULATIONS IN THE PHYSICAL SCIENCES

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193093 Published online by Cambridge University Press

Computer Simulations[1]

Paul Humphreys

University of Virginia

1. Introduction

A great deal of attention has been paid by philosophers to the use of computers in the modelling of human cognitive capacities and in the construction of intelligent artifacts. This emphasis has tended to obscure the fact that most of the high-level computing power in science is deployed in what appears to be a much less exciting activity: solving equations. This apparently mundane set of applications reflects the historical origins of modern computing, in the sense that most of the early computers in Britain and the U.S. were devices built to numerically attack mathematical problems that were hard, if not impossible, to solve non-numerically, especially in the areas of ballistics and fluid dynamics. The latter area was especially important for the development of atomic weapons at Los Alamos, and it is still true that a large portion of the supercomputing capacity of the United States is concentrated at weapons development laboratories such as Los Alamos and Lawrence Livermore.

Computer simulations now play a central role in the development of many physical sciences. In astronomy, in physics, in quantum chemistry, in meteorology, in geophysics, in oceanography, in crash analysis of automobiles, in the design of computer chips, in the planning of the next generation of supercomputers, in the discovery of synthetic pharmaceutical drugs, and in many other areas, simulations have become a standard part of scientific practice. My aim in the present paper is simply to provide a general picture of what computer simulations are, to explain why they have become an essential part of contemporary scientific methodology, and to argue that their use requires a new conception of the relation between theoretical models and their applications.[2]

Why should philosophers of science be interested in this new tool? Mostly, I think, because the way that simulations are developed and implemented forces us to reexamine a lot of what we tend to take as the right way to characterize parts of mathematically-oriented methodology and theorizing. Where this reexamination takes us will become clear as we go along, but before I discuss computer simulations specifically, I want to make some general points about the role of mathematical models in physical science. Let's begin with a claim that ought to be uncontroversial, but is not given enough emphasis in philosophy of science. The claim is: *One of the primary features*

PSA 1990, Volume 2, pp. 497-506

that drives scientific progress is the development of tractable mathematics. Whenever you have a sudden increase in useable mathematics, there will be a concomitant sudden increase in scientific progress in the area affected. This should not really need to be pointed out, but so much emphasis is placed on conceptual changes in science that powerful instrumental changes tend to be downplayed. This kind of sudden increase in mathematical power happened with the invention of the differential and integral calculus in the middle of the seventeenth century; it happened with the sudden explosion of statistical methods at the end of the nineteenth century, and I claim that the ability to implement numerical methods on computers is, in the late twentieth century, as significant a development as those earlier inventions. But what kind of development is it? Has it introduced a distinctively different kind of method into science, as Rohrlich (1991), for example, claims, or is it simply a technologically enhanced extension of methods that have long existed? If computer simulation methods are simply numerical methods, but greatly broadened in scope by fast digital computation devices with large memory capacity, then the second 'just much more of the same' view would be correct, and the situation would be similar to that in mathematics, where the introduction of computer-assisted proofs, such as were used to execute the massive combinatorial drudgery involved in the proof of the four colour theorem, is often regarded as not having changed the fundamental conception of what counts as a proof. My own view is that the situation is more complex than this simple dichotomy represents, because the introduction of computer simulation methods is not a single innovation but a multifaceted development. Let's begin with a couple of simple examples to show why mathematical intractability is an important constraint on scientific models.

2. Practical and Theoretical Unsolvability of Models

Take what is arguably the most famous law of all, Newton's Second Law. This can be stated in a variety of ways, but its standard characterization is that of a second order ordinary differential equation:

$$F = md^2y/dt^2 \qquad <1>$$

To employ this we need to specify a particular force function. In the first instance, take

$$F = GMm/R^2 \qquad <2>$$

as the gravitational force acting on a body near the Earth's surface (M is the mass of the Earth, R its radius). Then

$$GMm/R^2 = md^2y/dt^2 \qquad <3>$$

is easily solved. But the idealizations that underlie this simple mathematical model make it hopelessly unrealistic. So let's make it a little more realistic by representing the gravitational force as $GMm/(R + y)^2$, where y is the distance of the body from the Earth's surface, and by introducing a velocity-dependent drag force due to air resistance. We obtain

$$GMm/ (R + y)^2 - c\rho s(dy/dt)^2 - md^2y/dt^2 \qquad <4>$$

Suppose we want to make a prediction of the position of this body at a given time, supposing zero initial velocity and initial position $y = y_0$. To get that prediction you have to solve <4>. But <4> has no known analytic solution — the move from <3> to <4> has converted a second-order, linear, homogeneous ODE into a second-order, non-linear, homogeneous ODE, and the move from linearity to non-linearity turns

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193093 Published online by Cambridge University Press

simple mathematics into intractable mathematics. Exactly similar problems arise in quantum mechanics from the use of Schrodinger's equation, where different specifications for the Hamiltonian in the schema

$$H\Psi = E\Psi$$

lead to wide variations in the degree of solvability of the equation. For example, the calculations needed to make quantum mechanical, rather than classical, predictions in chemistry about even very simple reactions, such as the formation of hydrogen molecules when spin and vibration variables are included, are extremely difficult and have only recently been carried out. (An explicit discussion of the differences between ab initio and semi-empirical methods in quantum chemistry is given below.)

You might say that this feature of unsolvability is a merely practical matter, and that as philosophers we should be concerned with what is possible in principle, not with what can be done in practice. But recent investigations into decision problems for differential equations have demonstrated that for many algebraic differential equations [ADE's] (i.e. those of the form

$$P(x,y_1,...,y_n,y_1^{(1)},...,y_m^{(1)},...,y_1^{(n)},...,y_m^{(n)}= 0$$

where P is a polynomial in all its variables with rational coefficients) it is undecidable whether they have solutions. For example, Jaskowski (1954) showed that there is no algorithm for determining whether a system of ADE's in several dependent variables has a solution in [0,1]. Denef and Lipshitz (1984) show that it is undecidable whether there exist analytic solutions for such ADE's in several dependent variables around a local value of x. (Further results along these lines, with references, can be found in Denef and Lipshitz (1989)). Obviously, we cannot take decidability as a necessary condition for a theory to count as scientifically useful, otherwise we would lose most of our useful fragments of mathematics, but these results do show that there are in principle, as well as practical, restrictions on what we can know to be solvable in physical theories.[3]

There is a methodological point here that needs emphasis. While much of philosophy of science is concerned with what can be done in principle, for the issue of scientific progress what is important is what can be done in practice at any given stage of scientific development. That is, because scientific progress involves a temporally ordered sequence of stages, one of the things that influences that progress is that what is possible in practice at one stage was not possible in practice at an earlier stage. If one focusses on what is possible in principle (i.e. possible in principle according to some absolute standard, rather than relative to constraints that are themselves temporally dependent) this difference cannot be represented, because the possibility-in-principle exists at both stages of development. So although what is computable in principle is important for, say, the issue of whether computational theories of the mind are too limited a representation of mental processes, what is computable in practice is the principal feature of interest for the methodologies we are considering here.

This inability to obtain specific predictions from mathematical models is a very common phenomenon, because most non-linear ODE's and almost all PDE's have no known analytic solution. In population biology, for example, consider the Lotka-Volterra equations (first formulated in 1925)

$$dx/dt = ax + bxy$$

$$dy/dt = cy + dxy$$

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193093 Published online by Cambridge University Press

where x = population of prey, y = population of predators, a (>o) is the difference between natural birth and death rates for the prey, b(<o), d(>o) are constants related to chance encounters between prey and predator, c(<o) gives the natural decline in predators when no prey are available. With initial conditions x(o) = e, y(o) = f , there is no known analytic solution to the equation set.

These examples could be multiplied indefinitely, but I hope the point is clear: clean, abstract, presentations of theoretical schemas disguise the fact that the vast majority of those schemas are practically inapplicable in any direct way to even quite simple physical systems. This is not the point that models are never applicable to real systems: the point here is that even with radical idealizations, the problem of intractability is often inescapable, i.e. in order to arrive at an analytically treatable model of the system, the idealizations required would often destroy the structural features that make the model a model of that system type. This problem is widespread, and cuts across both sciences and subfields of those sciences, although it is more prevalent in some fields than in others.

These problems put severe limits on the applicability in practice of the standard, syntactically formulated method of hypothetico-deductivism, for most of the equations that represent the fundamental or derived theories of physics, chemistry, and so on cannot be used in practice to make precise deductive predictions from those representations together with the appropriate initial or boundary conditions. I should say here that I want to remain neutral as far as possible about the relative merits of the syntactic and semantic (or structuralist) reconstructions of theories. Although the semantic approach has definite advantages, both accounts are logical reconstructions of scientific practice. Because we are concerned here to stay as close as possible to considerations that present immediate problems to actual scientific practice, the debate over the merits of these reconstructions has only an indirect relevance to our interests. It is worth noting, however, that the issue of practical unsolvability means that the formulation of a theoretical model in some specific mathematical representation, rather than as a set of metamathematical structures, is an inescapable concern, and that whereas the semantic approach generally considers different linguistic formulations as mere linguistic variants of an underlying common structure, linguistic reformulations frequently have a direct impact on the ease of solvability of a mathematical representation, and hence this level cannot be ignored completely. In particular, I want to urge that what is of primary interest here is the mathematical form of equation types and not their logical form. To be specific: one could reformulate <1>, <2>, <3> and <4> in a standard logical language by using variable-binding operators, thus forcing them into the standard quantified conditional form that serves as the representation of laws in the traditional syntactic approaches, but to do this would be to distort what is crucial to issues of solvability, which is the original mathematical form.

It is this predominance of mathematically intractable models that is the primary reason why computational physics (and similar methods in other sciences), which provides a practical means of implementing non-analytic methods, constitutes a significant and, I think, a permanent, addition to the mathematical methodology of science.

3. Definitions of Computer Simulation

Here, taken more or less at random, are some suggestions that have been made for characterizing computer simulations: "Simulation is the technique by which understanding the behaviour of a physical system is obtained by making measurements or observations of the behaviour of a model representing that system." (Ord-Smith

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193093 Published online by Cambridge University Press

(1975), p.3) "This is what simulation is all about, i.e. experimenting with models" (ibid, p.3)

> A precise definition of simulation is difficult to obtain...the term simulation will be used to describe the process of formulating a suitable mathematical model of a system, the development of a computer program to solve the equations of the model and operation of the computer to determine values for system variables . (Bennet (1974), p.2)

> The mathematical/logical models which are not easily amenable to conventional analytic or numeric solutions form a subset of models generally known as simulation models. A given problem defined by a mathematical/logical model can have a feasible solution, satisfactory solution, optimum solution or no solution at all. Computer modelling and simulation studies are primarily directed towards finding satisfactory solutions to practical problems. (Neelamkavil (1987), p.1).

> Simulation is a tool that is used to study the behaviour of complex systems which are mathematically intractable. (Reddy (1987), p.162)

Because of the variety of uses to which the term 'simulation' has been put, I am reluctant to try to formulate a general definition. It would be more profitable at this stage to simply explore the methods that are used under categories 1), 2), and 3) in section 4 below. We can, however, formulate a working definition based on the last characterization, which needs to be modified in three ways. First, simulation is a set of techniques, rather than a single tool. As the other quotations indicate, it would be hard to make a case for the view that there is an underlying unity to the set, at least at the present state of development of the field. Second, the systems that are the subject of simulations need not be complex either in structure or behaviour. As we also saw earlier, mathematical intractability can affect differential or integral equations having a quite simple mathematical structure, as in the case of the motion of the body falling under the influence of gravity, subject to a velocity-dependent drag force. The behaviour of this system is not unduly complex, merely hard to predict quantitatively without numerical techniques. Third, many computer simulations turn analytically intractable problems into ones that are computationally tractable, and we do not want to exclude numerical methods as a part of mathematics.

We thus arrive at the following working definition which captures what is common to almost all the simulations with which I am familiar.

Working Definition. A computer simulation is any computer-implemented method for exploring the properties of mathematical models where analytic methods are unavailable.

Some further remarks may be helpful. Although the everyday use of the term 'simulation' has connotations of deception, so that a simulation has elements of falsity, this has to be taken in a particular way for computer simulations. Inasmuch as the simulation has abstracted from the material content of the system being simulated, has employed various simplifications in the model, and uses only the mathematical form, it obviously and trivially differs from the 'real thing', but in this respect, there is no difference between simulations and any other kind of mathematical model, and it is primarily when computer simulations are used in place of empirical experiments that this element of falsity is important. But if the underlying mathematical model can be realistically construed (i.e. it is not a mere heuristic device) and is well-confirmed,

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193093 Published online by Cambridge University Press

then the simulation will be as 'realistic' as any theoretical representation is. Of course, approximations and idealizations are often used in the simulation that are additional to those used in the underlying model, but this is a difference in degree rather than in kind.

Next, in order for something to be a computer simulation, the whole process between data input and output must be run on a computer, whereas computational physics can involve only some stages in that process, with the others being done 'by hand'. Third, because computer simulations are usually oriented towards approximate solutions rather than exact solutions, they can be viewed as optimization devices that sometimes involve satisficing criteria. This approach underlies the variational method mentioned earlier, it underlies the simulated annealing method frequently used in connectionist models of perception and problem solution (see McClelland and Rumelhart (1986), especially Chapter 6), and it underlies many other intuitive 'good enough' criteria used in other areas.

4. Can Computer Simulation Be Identified With Numerical Methods?

What is computer simulation? The terminology is so widely used that it is hard to find a core meaning, but here are some central uses:

1) To provide solution methods for mathematical models where analytical methods are presently unavailable.
2) To provide numerical experiments in situations where natural experimentation is inappropriate (for practical reasons) or unattainable (for physical reasons). Under the former lie experiments that are too costly, too uncertain in their outcome, or too time consuming. Under the latter lie such experiments as the rotation of angle of sight of galaxies, the formation of thin disks around black holes, and so forth.
3) To generate and explore theoretical models of natural phenomena.

It may seem that use 1) is simply the use of numerical methods for solution purposes. To examine this claim, we need some definitions. *Numerical mathematics* is concerned with obtaining numerical values of the solutions to a given mathematical problem. *Numerical methods* is the part of numerical mathematics concerned with finding an approximate, feasible, solution. *Numerical analysis* has as its principal task the theoretical analysis of numerical methods and the computed solutions, with particular emphasis on the error between the computed solution and the exact solution.

Can we identify numerical methods with computer simulations? Not directly, because there are at least two additional features that a numerical method must have if it is to count as a computer simulation. First, the numerical method must be applied to a specific scientific problem in order to be part of a computational simulation. Second, the method must be computable in real time and be actually implemented on a concrete machine.

Beyond this, there is an important potential distinction between uses 1) and 3). In 1), the development of the model is made along traditional lines: some more or less fundamental theory is brought to bear on the phenomenon, theory which at least in its abstract, general, form is well understood and confirmed. Deductive consequences are drawn out from this theory to bring the general theory into contact with the specific area under investigation, and then the computational implementation of these consequences constitutes the simulation of the system. In contrast, in use 3), the development of the models is partly empirical, partly theoretical, and partly heuristical, with

exploration and feedback from the simulation playing an important role in this development.

This distinction is not clearcut, and especially in use 3), elements from uses 1) and 2) often play a significant role. The difference is similar to a distinction that is often drawn in quantum chemistry between *ab initio* methods and semi-empirical methods (see e.g. R. McWeeny and B.T.Sutcliffe (1969), Chapter 9). Three kinds of treatments can be used to predict the energy levels of molecular orbitals. *Ab initio* methods use the actual Hamiltonian for the system in Schrodinger's equation. Idealizations are made, such as a fixed nucleus, only electrostatic interactions between particles, and non-relativistic calculations, and these idealizations are often drastic, but the goal is to represent as many of the important features of the molecules as possible. Then using a 'trial function' it calculates the solution 'exactly'. Semiempirical methods estimate some parameters in the orbital states that are difficult to calculate directly by empirical data or by numerical approximation, and then proceed as in the ab initio case. Model level methods use a Hamiltonian that deliberately omits some important influences on the energy levels, such as inter-electron interactions.

The distinction here between *ab initio* methods and model level methods seems to me to be quite arbitrary, since both use idealizations, and the interesting difference is that between *ab initio* methods and semi-empirical methods, and this is an appropriate place to discuss the differences between fundamental and phenomenological models. This distinction reflects the 'bottom up' and 'top down' methods familar from other areas of methodology, and there is a significant divergence of views about whether models should be constructed on the basis of some underlying general theoretical considerations, or whether instrumentally successful but theoretically ungrounded models should be used when the theoretical approach is infeasible. Both kinds are used in simulations and I see no reason to deny the appropriateness of either. I choose to focus on fundamental models here, primarily for two reasons. The first is one of expertise, or lack of it. Phenomenological models are usually highly specific devices constructed for the purpose of representing some specific phenomenon. A great deal of physical, chemical, or biological knowledge goes into their assessment, justification, and use (this is one area where 'physical intuition' is clearly an important consideration) and for this reason, such simulations can be assessed only by those actively working with them. The second reason for emphasizing fundamental models here is that this makes a comparison of simulation methodology with traditional philosophical views on theory structure and application much easier, for the latter is oriented almost exclusively towards fundamental theory. This said, a few remarks about the relation between the two approaches in the case of *ab initio* and semi-empirical models might be appropriate. (Semi-empirical models are not the same as phenomenological models, in that the former are still guided to a considerable extent by theory, but for the first reason just mentioned, I am not in a position to address phenomenological models in any detail.)

One important result of the availability of large-scale computational power is that whereas many idealizations in models, or the use of semi-empirical methods, were once forced upon chemists because the model had to result in tractable analytic mathematics, the idealizations made in *ab initio* methods now need not be determined primarily by that constraint, but are set by (a) limits on computational power available, (b) the ability to mathematically represent in the Hamiltonian complex influences on the energy levels (c) the availablity of numerical methods to approximate the representations in (b). This is a clear example of computational chemistry: the use of computers to allow one to treat models that could not be used without them. Indeed, these methods illustrate an interesting trade-off: These numerical methods allow one to deal with more realistic theories, and the increased use of approximations in the math-

ematics allows a decreased use of idealizations in the physics. This still leaves the treatment of most molecules currently outside the scope of *ab initio* methods, and given the restrictions due to (a) that are discussed below, no purely *ab initio* method will ever be fully computationally feasible, but the important point is that more and more systems that were once untreatable by fundamental approaches can now have theoretically justified quantum mechanical methods brought to bear on them. Compare this with methods that rely on the variation theorem. (See Eyring et al (1944) for a development of this theorem). The theorem states "If a normalized trial function S satisfies the relevant boundary conditions but is otherwise arbitrary, then $\langle S/H/S \rangle \geq E_0$, where E_0 is the lowest eigenvalue, the equality applying when S is an exact solution." (Further applications of this procedure can be used to find approximations to other eigenfunctions.) Then the best wave function is obtained by varying the parameters in a trial function until the lowest energy is obtained. Here, the computational methods allow exploratory investigations that would not be possible without computers, and these are different from theoretically based methods in that although theory may be used as a guide to which parametric family of functions to explore (Gaussian or Slater orbitals are usually used, however, making the contribution of theory minimal), the final result is a matter of computational trial and error rather than explicit theoretical derivation. Moreover "A wave function that gives a good [estimate of the] energy does not necessarily give a particularly good value for another quantity, for example the dipole moment, whose expectation value may arise principally from somewhat different regions of space." (McWeeny & Sutcliffe, op. cit., p.235). I also find these figures from McWeeny and Sutcliffe, op.cit, p.239 revealing: Abstract developments of quantum mechanics require an infinite set of basis vectors to represent states. For the finite basis sets that actual applications need, suppose that m atomic orbitals are used (in the linear combination of atomic orbitals representation of molecular orbitals — the LCAO method). Then one needs $p = m(m+1)/2$ distinct integrals to calculate one-electron Hamiltonians, and $q = p(p+1)/2$ distinct integrals to calculate electron interaction terms. This gives

m =	4	10	20	40
q =	55	1540	22155	336610

This is a clear case where computational constraints, which are extra-theoretical and here involve primarily memory capacity, place severe limitations on what can be done at any given stage of technological development. This is different in principle, I think, from the constraints that the older analytic methods put on model development, because there new mathematical techniques had to be developed to allow more complex models, whereas in many cases in computational science, the mathematics stays the same, and it is technology that has to develop. The use of trial orbitals that I mentioned earlier in connection with the variational method seems to show that a very crude model can give an apparently realistic representation of the system. That is, in deciding upon the appropriate potential energy function to use in the Hamiltonian, suppose we choose one corresponding to a Slater atomic orbital of the form

$$V(r) = -fn/r + [n(n-1) - l(l+1)]/2r^2$$

where f is a parameter representing the effective field affecting the electron. Then using these atomic orbitals as the finite basis, we have to decide where the expansion of the state function will be truncated. Then, given various trials R, and trial orbitals S for another electron, we have to minimize the energy

$$E = 2\langle R/h/R \rangle + (2\langle RS/g/RS \rangle - \langle RS/g/SR \rangle)$$

https://doi.org/10.1086/psaprocbienmeetp.1990.2.193093 Published online by Cambridge University Press

Although there is a great deal of computation involved here, and certainly trial and error 'experimentation', there is also a good deal of theory lying in the background to justify the method, and even though the atomic orbitals used are pretty crude approximations, they are still guided by a physical model that has some theoretical justification.

I am thus going to treat each of uses 1), 2) and 3) above as part of computational physics (chemistry, etc.), and to consider computer simulation as a subset of the methods of computational science. Much more needs to be said about what is special to simulations, but I hope that the example just discussed shows that the interplay between theory, experiment, and computation in computational science entails that it is not to be identified with numerical methods, and a fortiori, neither should computer simulations.

Notes

[1]Research for this paper was supported by NSF grant DIR-8911393. I should like to thank Fritz Rohrlich for helpful discussions in connection with the PSA symposium.

[2]When examining this activity, we must be wary of one thing, which is that the field of computer simulation methods is relatively new and as such is rapidly evolving. Techniques that are widely used now may well be of minor interest twenty years hence, as developments in computer architecture, numerical methods, and software routines take place. The specific details of different kinds of simulation methods, such as finite-difference methods and Monte Carlo methods will be explored in a future paper,and some examples of currently used simulations are given in the following paper by Rohrlich.

[3]A further source of difficulty, at least in classical mechanics, involves the imposition of nonholomorphic constraints (i.e. constraints on the motion that cannot be represented in the form

$$f(r_1,...,r_n,t) = 0$$

where $\{r_i\}$ are the spatial coordinates of the particles comprising the system). For a discussion of these constraints, see Goldstein (1980), pp.11-14.

References

Bennett, A.W. (1974), *Introduction to Computer Simulation.* St. Paul, West Publishing Co.

Denef, J. and Lipshitz, L. (1984), "Power Series Solutions of Algebraic Differential Equations", *Mathematische Annalen*, 267:213-238.

_____. (1989), "Decision Problems for Differential Equations" *J. Symbolic Logic* 54:941-950.

Eyring, H., Walter, J., and Kimball, G. (1944), *Quantum Chemistry*. New York, J. Wiley and Sons.

Goldstein, H. (1980), *Classical Mechanics*, (2nd Edition). Reading, Mass, Addison-Wesley Publishing Co.

Jaskowski, S. (1954), "Example of a Class of Systems of Ordinary Differential Equations Having No Decision Method for Existence Problems" *Bulletin de l'Academie Polonaise des Sciences*, Classes III, vol. 2, pp. 155-157.

McClelland, J. and Rumelhart, D. (1986), *Parallel Distributed Processing*, Vols 1,2. Cambridge, The MIT Press.

McWeeny, R. and Sutcliffe, B.T. (1969), *Methods of Molecular Quantum Mechanics*, New York, Academic Press.

Neelamkavil, F., (1987), *Computer Simulation and Modelling*, New York, J. Wiley.

Ord-Smith, R.J. and Stephenson, J. (1975), *Computer Simulation of Continuous Systems*. New York, Cambridge University Press.

Reddy, R. (1987), "Epistemology of Knowledge-Based Systems", *Simulation*, 48:161-170.

Rohrlich, F. (1991), "Computer Simulation in the Physical Sciences", in PSA 1990, Volume 2, A. Fine, M. Forbes, and L. Wessels (eds). East Lansing: Philosophy of Science Association, pp. 497-506.