Chapter 2
Simulation and Artificial Life

2.1 Overview

Before beginning this extensive analysis of agent-based modelling, the philosophical and methodological background of simulation science must be discussed. Scientific modelling is far from new; conceptual and mathematical models have driven research across many fields for centuries. However, the advent of computational modelling techniques has created a new set of challenges for theorists as they seek to describe the advantages and limitations of this approach.

After a brief discussion of the historical foundations of scientific modelling, both mathematical and computational, some distinctions that characterise modelling endeavours will be described. The distinction between models for science and engineering problems provides a useful framework in which to discuss both the goals of modelling and the difficult problem of validating models. These discussions are framed in the context of artificial life, a field which depends upon computational modelling as a central methodology.

This chapter lays the groundwork for the next two chapters to come, and by extension Part I of this text as a whole. The discussion here of emergence, and the related methodologies of 'bottom-up' modelling, will allow us to understand the major philosophical issues apparent in the field of computational modelling. This, in turn, will prepare us for an in-depth discussion of the field of Artificial Life in the following chapter, which will provide a backdrop for the discussion of general issues in biological modelling in Chap. 4. Similarly, the discussion here of the debate regarding the explanatory capacity of simulation models will be a recurring thread throughout this text, in all three parts of the argument.

E. Silverman, *Methodological Investigations in Agent-Based Modelling*,
Methodos Series 13, https://doi.org/10.1007/978-3-319-72408-9_2

2.2 Introduction to Simulation Methodology

2.2.1 The Goals of Scientific Modelling

The construction of models of natural phenomena has been a continuous feature of human scientific endeavour. In order to understand the behaviour of the systems we observe around us, we choose to construct models to allow us to describe that behaviour in a simplified form. To use our central bird migration example, imagine that a researcher wishes to describe the general pattern of a specific bird species' yearly migration. That researcher may choose to sit and observe the migrations of that species year-on-year for a lengthy period of time, allowing for the development of a database showing that species movement over time.

However, a model of that species' migration patterns could save our researcher a great deal of empirical data collection. A model which takes the collected empirical data and derives from it a description of the bird species' general behaviour during each migration season could provide a means of prediction outside of a constant real-world observation of that species. Further, one can easily imagine models of the migration problem which allow for detailed representation of the birds' environment and behaviour, allowing for a potentially deeper understanding of the causes of these migrations.

Thus, in the context of this discussion, the goals of scientific modelling encompass both description and explanation. A simplified description of a phenomenon is inherently useful, reducing the researcher's dependence on a continuous flow of collected empirical data, but explanation is the larger and more complex goal. Explanation implies understanding: a cohesive view of the forces and factors that drive the behaviour of a system. To develop that level of understanding, we must first understand the nature and limitations of the tools we choose to employ.

2.2.2 Mathematical Models

Throughout the history of science, mathematical models have been a vital part of developing theories and explanations of natural phenomena. From Newton's laws of motion to Einstein's general relativity, mathematical models have provided explicit treatments of the natural laws that govern the world around us. As we shall come to understand, however, these models are particularly well-suited to certain classes of phenomena; the concise description of natural laws that follows from a mathematical model becomes ever more difficult to develop as the phenomena in question becomes more complex.

Even with the appealing simplicity of a mathematical description of a phenomenon however, certain methodological difficulties come into play. Some models may require a very complex system of linked equations to describe a system, creating a vast number of parameters which are not known a priori (see Chap. 4 for discussion of this in relation to population biology). These black-box simulations create difficulties for both theorists and experimentalists; theorists struggle to set values for those parameters, while experimentalists likewise struggle to validate models with such a range of potential parameters.

There is also the question of the accuracy of a given mathematical model, which can be difficult to determine without repeated testing of that model's predictive capacity. For example, the motion of objects in space can be described by Newton's laws of motion; however, for very large bodies which are affected by the tug of gravity, we must use Einstein's general relativity to describe that motion. If we turn the other way and wish to describe the motions of atoms and particles, then we must use quantum mechanics. The intersections of those models, particularly that of Einstein's relativity and quantum mechanics, are far from easy to develop; the fabled union of these two theories has occupied physicists for decades and will likely continue to do so for some time (Gribbin 1992).

2.2.3 Computational Models

The advent of easily-available computing power has revolutionised the process of scientific modelling. Previously intractable mathematical models have been tractable through the sheer brute-force calculating power of today's supercomputers. Supercomputers now allow physicists to run immense n-body simulations of interacting celestial bodies (Cox and Loeb 2007), model the formation of black-hole event horizons (Brugmann et al. 2004) and develop complex models of global climate and geophysics (Gregory et al. 2005).

Beyond the sheer number-crunching power of computational methods, the flexibility of computational modelling has resulted in the development of new varieties of models. While the specific characteristics of each simulated celestial body in a model of colliding galaxies is relatively unimportant, given that each of those galaxies contains billions of massive bodies with similar gravitational impacts on surrounding bodies, certain other phenomena depend on complex individual variation to be modelled accurately. Evolution, for example, depends upon the development of new species through processes of individual mutation and variation mediated by natural selection (Darwin 1859); thus, a model of evolving populations requires a description of that individual variation to be effective.

Take once again our central example. If our hypothetical bird-migration researcher hypothesizes that migration behaviour is due to evolutionary factors that he may be able to represent in a model, then he must be able to represent the effects of biological evolution in some form within that model. If he wishes to represent those effects, the digital birds within his model would require individual complexity that can produce variation within the simulated species. In contrast, if he were modelling only the patterns of the bird movements themselves, then he need only represent the impact of those movements on the other agents in the simulation, as in the colliding galaxy model above; individual variation in those agents is not such a driving force behind the patterns of bird movements as it would be in an evolutionary model of the development of those movements.

2.2.4 The Science Versus Engineering Distinction

As computational modelling has developed, so has a distinction between varying computational approaches. Some models seek to provide predictive power related to a specific physical or natural phenomenon, while others seek to test scientific hypotheses related to specific theories. The first type has been characterised as modelling for engineering, and the second has been described as modelling for science (Di Paolo et al. 2000; Law and Kelton 2000).

Models for engineering are dependent on empirical data to produce predictions. For example, a transportation engineer may wish to examine the most efficient means for setting traffic signals (Maher 2007). To determine this, the modeller will examine current signal settings, note the average arrival time of vehicles at each junction, analyse the anticipated demand for each service, and other similar factors. With this information in hand the engineer can produce a model of the current operating traffic pathways, and alter parameters of those simulated services to attempt to produce an optimum scheduling algorithm for the new signals.

Similarly, to stretch our bird example to the engineering realm, imagine that our migration researcher has decided to model a the dissemination of information via messenger pigeon. If he wishes to find an optimum schedule on which to release and retrieve these pigeons, he could use an engineering-type model to solve this problem. He could examine current and past messenger-pigeon services, note the average transit time for the delivery and retrieval of those messages, and the level of rest and recuperation needed by each bird. With an understanding of these factors, the researcher could develop a model which would provide optimum release schedules, given different potential numbers of birds and varying demand for message delivery.

Models for science, in contrast, focus instead on explanation and hypothesis-testing. A model of the development of animal signalling by its very nature cannot depend on the availability of empirical data; after all, we cannot simply watch random evolving populations in the hope that one of them may develop signalling behaviours while we wait. Instead, the model is based upon a hypothesis regarding the contributing factors that may produce the development of signalling behaviours; if the model produces a simulated population which displays those behaviours, then the modeller may attribute more validity to that hypothesis. This approach is the focus of this text.

Returning to the bird example, our researcher would be taking a similarly scientific modelling perspective if he wished to construct a model which illustrates the influence of individual bird movements on migrating flocks. He hypothesizes that individual birds within the flock assume controlling roles to drive the timeliness of the flock's migratory behaviour. He could test such a hypothesis by developing a simulated population of migrating birds in which certain individual agents within the simulation can affect the movement of multiple other migrating agents; if the presence of those controlling agents appears to confirm the necessity of such individuals to keep migrations moving in an appropriate timeframe, then he may propose that such mechanisms are important to real-world migratory behaviour.

2.2.5 Connectionism: Scientific Modelling in Psychology

The advent of this sort of scientific modelling has produced not only new types of models, but new fields of enquiry within established fields. The development of the connectionist approach to the study of behaviour and cognition provides one example of the scientific modelling perspective, and introduces us to the concept of emergent explanations.

The development of computational modelling techniques together with advances in neuroscience led some researchers to investigate models of neural function. These neural network models consist of simplified neuronal units, with specified activation thresholds and means of strengthening or weakening synaptic connections, which aim to reproduce the neural basis of behaviours (Rumelhart and McClelland 1986). The idea that models of this nature could demonstrate the emergence of cognitive behaviour brought with it related ideas concerning the mind that caused some controversy.

In order for the connectionist to assert that their model can represent cognitive behaviour, one must assume that mental states correspond to states of activation and connection strengths in a given neural network. This concept was derided by some who viewed this as an overly reductionist stance, and that in fact the symbolic manipulation capability of the mind is crucial to understanding cognitive function (Fodor and Pylyshyn 1988). This is related to the perspective espoused by many in the field of artificial intelligence, in which the manipulation of symbols was considered essential to the development of intelligence (Newell and Simon 1976).

The connectionist thus forms scientific models which aim to test the hypothesis that learning mechanisms in neural networks can lead to the development of cognition. While psychologists are certain that the human brain functions via the interaction of billions of individual neurons, the degree of correspondence between these neural-network models and actual brain function is debatable (see Pinker and Mehler 1988 for a damning critique of the unrealistic results of a connectionist model of language, as one example). Many models of this type use idealised neuronal units to investigate possible explanations of behaviour, such as creating non-functional 'lesion' areas in a network designed to perform visual search tasks as a means of theorising about possible causes of visual deficits (Humphreys et al. 1992). In this respect, these sorts of connectionist models are designed to test hypotheses and develop theories rather than generate predictions based on empirical data.

2.2.6 Bottom-Up Modelling and Emergence

As noted above, the controversy surrounding connectionist modelling hinged upon one of its base assumptions: the idea that the low-level interaction of individual neuronal units could produce high-level complex behaviour. The cognitivists and computationalists of psychology found this distressing, as reducing cognition to collections of neural activations eliminates the concept of symbolic manipulation as a precursor to thought and cognition; the 'distributed representation' concept proposed by connectionists would remove the necessity for such higher-level concepts of discrete symbol manipulation by the brain (Fodor and Pylyshyn 1988).

Of course, such perspectives need not necessarily be diametrically opposed. One can certainly imagine connectionism forming a useful element of the study of cognition, with the cognitivists continuing the study of mental representation and symbol manipulation in relation to larger concepts of mental behaviour that are less well-suited to the connectionist modelling perspective. Indeed, given that connectionist systems can implement symbol-manipulation systems, these philosophical differences seem minor (Rowlands 1994).

However, the idea of this type of 'bottom-up' modelling is crucial, and as a consequence the debate over this type of modelling bears great relevance to our discussion. The view proffered by connectionists that models of low-level interacting units can produce the emergence of higher-level complexity is a central element of the modelling perspectives being analysed in this text. The particular relevance of this controversy when considering problems of validation and scientific explanation will be examined further both in this chapter and in Chaps. 6 and 7 in particular.

2.3 Evolutionary Simulation Models and Artificial Life

2.3.1 Genetic Algorithms and Genetic Programming

Connectionism was far from the only prominent example of a computational innovation taking cues from biology. The use of genetic algorithms, modelled on the processes of biological evolution, has a long history in the computational sciences. Given the extensive study of natural selection in biological systems as an optimisation process, and the need for increasingly innovative optimisation techniques within computer science, the use of an analogue of that process for computational applications seemed a natural fit. Indeed, as early as the 1950s the computers available were being put to use on just these sorts of problems (Fraser 1957).

Since these early days of experimentation, the genetic algorithm became an established method of optimisation in certain problem spaces. Such algorithms seek to encode potential solutions to the problem at hand in forms analogous to a biological 'genotype'; each solution is then examined to determine its suitability as a solution, and evaluated according to a specified fitness function. The most fit solutions can then be combined, individual mutations can be generated if desired, and the next generation of potential solutions is subjected to the same process. Over many generations, the genetic algorithm may find a solution suitable for the problem, and in some cases the solution comes in an unexpected and novel form due to the influence of this selection pressure. Such systems came into the spotlight quite prominently in the 1970s, when John Holland's book on the topic provided a strong introduction (Holland 1975); later works by Goldberg (1989), Fogel (1988), and Mitchell (1996) cemented the position of genetic algorithms as a useful method for optimisation and search problems.

Genetic algorithms do suffer from methodological difficulties, of course; certain problems are not well-suited to genetic algorithms as a means of finding appropriate solutions. In addition, the design of a useful fitness function can be extremely difficult, as the programmer must be careful to avoid solutions which cluster around local optima (Mitchell 1996). Incorporating an appropriate amount of variation in the generated population can be vital for certain applications as well, as the right level of random mutation can provide a useful means to escape those local optima.

2.3.2 Evolutionary Simulations and Artificial Life

While genetic algorithms became popular amongst certain elements of the computer science community, they also drew great interest from those interested in the biological function of evolution. As the artificial intelligence community sought to model the fundamentals of human intelligence and cognition, others sought to use computational methods to examine the fundamentals of life itself.

The field of artificial life, or ALife, has complex beginnings, but is most often attributed to Langton (2006) who first christened the field with this title. ALife however has strong links with the artificial intelligence community (Brooks 1991), as well as with the earlier modelling traditions of ecology and population biology. The influence of the artificial intelligence community, the interest in bottom-up modelling as alluded to earlier in our review of connectionism, and the development of new techniques to produce adaptive behaviour in computational systems all seem to have had a hand in the development of ALife.

ALife work to date has revolved around a number of related themes, but all of them share some method of reproducing the mechanics of biological adaptation in computational form. Genetic algorithms as described above are perhaps the most prominent example, with a great number of evolutionary simulations using such algorithms or some version thereof to provide that element of adaptation. While the members of this growing research community moved forward with these methods of simulating evolutionary systems, a related set of new challenges faced that community.

2.3.3 Bedau and the Challenges Facing ALife

Mark Bedau's 2003 (Bedau 2003) summary of the field of artificial life provides one view of the array of potential challenges facing the ALife researcher. In Bedau's view, ALife clearly displays the potential to enhance our understanding of the processes of life, and numerous fundamental questions that spring from those processes:

How does life arise from the non-living?

1) Generate a molecular proto-organism in vitro.
2) Achieve the transition to life in an artificial chemistry in silico.
3) Determine whether fundamentally novel living organisations can arise from inanimate matter.
4) Simulate a unicellular organism over its entire lifecycle.
5) Explain how rules and symbols are generated from physical dynamics in living systems.

Bedau begins his extensive list of ALife challenges with a look at the potential for these new methods of simulation to simulate the origins of life. Of course there is great debate over the best means to simulate such early beginnings. Simulating the development of cell structures has been an important theme in ALife (e.g., Sasahara and Ikegami 2004), as well as the development of simple self-replicating structures (Langton 1990). This is not entirely surprising, given that Von Neumann's self-replicating cellular automaton was clearly an influence on those seeking to understand the development of such forms in silico (Von Neumann and Burks 1966).

However, at what point might we agree that such self-replicating digital organisms have achieved a 'transition to life' as proposed by Bedau? At what point does that simulated organism become an instantiation of the laws governing the development of natural life? Agreement here is hard to come by; some argue that the status of 'alive' is best conferred on organisms that can self-reproduce (see Luisi (1998) for an evaluation of this and other definitions), while others argue that self-motility[1] is a more important determining factor (Hiroki et al. 2007), and still others appeal to the concepts of self-organisation and autopoiesis[2] (Maturana and Varela 1973). This issue hinges upon the theoretical perspective of the modeller to a large degree: if one believes that the properties of life are just as easily realised in the digital substrate as they are in the biological substrate, then an ALife simulation can easily achieve life (given an appropriate definition of such) regardless of its inherent artificiality. The issue of artificiality in ALife research and its import for the theorist and experimentalist are explored in detail in Chap. 3.

[1]The ability to move spontaneously or non-reactively. This is considered a vital capability for biological life – self-motility allows living things to move in pursuit of food sources, for example. See Froese et al. (2014) for a detailed exploration.

[2]Autopoietic systems are systems that can produce and sustain themselves through their own internal processes, such as the biological cell. The concept was originally described in relation to biological systems, but has since been adapted to characterise cognitive and social systems as well.

What are the potentials and limits of living systems?

6) Determine what is inevitable in the open-ended evolution of life.
7) Determine minimal conditions for evolutionary transitions from specific to generic response systems.
8) Create a formal framework for synthesizing dynamical hierarchies at all scales.
9) Determine the predictability of evolutionary manipulations of organisms and ecosystems.
10) Develop a theory of information processing, information flow, and information generation for evolving systems.

Bedau's next set of challenges are reminiscent of one of Chris Langton's more famous descriptions of artificial life, in which he stated that ALife could seek to examine 'life-as-it-could-be' rather than simply 'life-as-we-know-it' (Langton 1992). In other words, given that the ALife researcher can construct an enormous variety of possible models, and thus living systems if we agree that life can be realised in silico, then ALife can be used as a platform to understand the vast variety of potential forms that life can create, rather than only examine the forms of life we currently perceive in the natural world.

Bedau is alluding to similar ideas, proposing that ALife researchers can use their work to probe the boundaries of the evolution of life. By simulating evolutionary systems, he posits that we may be able to investigate the mechanics of evolution itself in a way impossible in conventional biology. The researcher is able to freely tweak and direct the evolutionary processes at work in his simulation, and if we accept that the simulation adequately represents the function of evolution in the real world, then such research may allow for a greater understanding of the limits of the evolutionary process.

How is life related to mind, machines and culture?

11) Demonstrate the emergence of intelligence and mind in an artificial living system.
12) Evaluate the influence of machines on the next major evolutionary transition of life.
13) Provide a quantitative model of the interplay between cultural and biological evolution.
14) Establish ethical principles for artificial life. (Bedau 2003, p. 506)

Finally, Bedau closes his list of ALife 'grand challenges' with more speculative notions of relating the development of life with the development of mind and society. He posits that cognition originates from similar roots as life, in that such mental activity is a biological adaptation like any other seen in evolving systems, and that in this context artificial life may provide insight into the origins of mind as well as life.

The idea that mind and culture follow similar rules of adaptation to life is not a new one; the field of evolutionary psychology is well-established, if controversial (Buss 2004), and Dawkin's discussion of the 'meme' in relation to cultural evolution is one of the more prominent examples of such thinking in sociology (Dawkins 1995). The question of whether simulation can become sufficiently sophisticated to allow for the emergence of these higher-order phenomena is critical to our upcoming examination of simulation in the social sciences; in fact, the same philosophical difficulties that face modellers of cognition have been linked with similar difficulties in using simulation to model the roots of society and culture (see Sawyer 2002, 2003, 2004 and the accompanying discussion in Chap. 5).

2.4 Truth in Simulation: The Validation Problem

2.4.1 Validation and Verification in Simulation

While Bedau has provided a useful summary of the possible challenges facing artificial life in the research realm, numerous other challenges also loom in the area of methodology for ALife modellers. The difficulty of tying simulation results to the system being simulated is one not confined to ALife, but instead is common to all varieties of simulation endeavour.

The validation and verification of simulations is often most troublesome for modellers of all types. Once a model is designed and run, the researcher must be able to express confidence that the results of that model bear a direct relation to the system of interest. This is validation, and in relation to computational models specifically, Schlesinger's description of this process as a 'substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model' (Schlesinger et al. 1979) is often cited. In other words, the modeller must demonstrate that the model displays a measure of accuracy within the domain to which the model has been applied.

Going hand-in-hand with validation is the concept of verification. A verified model is one in which the construction of the model is known to be accurate according to the framework in which that model is designed (Law and Kelton 2000). In relation to computational models specifically, this means that the model must be programmed appropriately to produce results that reflect the intent of the model's construction; given the inherent complexity of the software design process, this is not necessarily a simple task to complete.

2.4.2 The Validation Process in Engineering Simulations

Validation in simulations for engineering purposes, as described earlier, would tend to follow a certain pattern of verifying assumptions and comparing model predictions to empirical data (Sargent 1982, 1985). To illustrate these concepts we may return to the bird migration example. If our migration researcher wished to construct a model which provides an illustration of the migration behaviour of a certain bird species, he would first need to discuss the assumptions inherent in the conceptual model leading to the simulation's construction. For example, the speed and direction of movement of his simulated migrating populations should match those values gleaned from empirical observation; in general, his conceptual model should be informed by the available data. This verification of the conceptual model provides confidence that the assumptions made to produce the model have a solid foundation in theory related to the migrations observed in that bird species.

Having verified the conceptual model, the researcher would then need to verify the model itself. While confidence in the assumptions in the conceptual model has been established, the researcher must confirm that these assumptions have been implemented appropriately in the model itself. Has the code for the simulation been written correctly? Are the correct parameters in place in the model, given the parameters required by the conceptual model? If the researcher can demonstrate that the simulation has been implemented correctly, then validation can proceed.

The validation step is the most difficult, requiring a comparison of model data with real data. With our model migrating birds in place, the simulation should provide predictions of the behaviour of those birds when it is run. Do these simulation runs provide data that correlates appropriately with observational studies of that bird species' migration? Does empirical data related to bird migrations in general imply that the model's results are believable? Such questions can be complicated to answer, but nevertheless can be answered with access to appropriate empirically-collected data (see Law and Kelton 2000 for more discussion).

2.4.3 Validation in Scientific Simulations: Concepts of Truth

The procedure outlined above is undoubtedly complex, but achievable for the engineer. For the scientist, however, the procedure becomes more complex still. In a simulation which is designed to test hypotheses, and in which a clear relation to empirical data is not always obvious are indeed possible, the prospect of validating the results of that simulation depends on different sorts of relations between theory, data and model.

Appealing to the philosophy of science, Alex Schmid describes three theories of truth for simulations in science: the correspondence theory, the consensus theory, and the coherence theory (Alex Schmid 2005). The correspondence theory holds that the simulation must correspond directly with facts in reality; the consensus theory holds that the simulation must be acceptable under idealised conditions; and the coherence theory holds that the simulation must form a part of a coherent set of theories related to the system of interest.

The correspondence theory of truth relates closely to the methods of validation discussed in relation to engineering: our example bird migration simulation must display results that correspond directly to empirical data regarding migrating birds, otherwise the predictions of that simulation are of little use. Such a view coincides with the prevailing views of validation present in engineering models: validated models in the engineering perspective must demonstrate a close relationship to known empirical data about the problem under study.

The consensus theory, however, is much less defined, depending as it does on a more communal evaluation of the truth of the model. Our bird migration model may not provide entirely accurate predictions for future migration behaviours, but under this view we may still consider the simulation to be validated if the model is generally illustrative of bird migration behaviour. A more general version of our migration simulation, one developed as an abstract model not tied to any particular bird species, could fall into this category.

The coherence theory moves even further from the concept of truth most applicable to engineering, requiring only that the model in question fit into a given set of coherent beliefs. If our bird migration model fits cohesively into the general body of animal behaviour theory relating to migrating populations, then the model may provide a useful and valid addition to that theory. However, as Schmid points out, there is no reason that a coherent system of beliefs cannot also be completely false (Alex Schmid 2005); a model of our migrating birds travelling across a flat planet may be accurate given the belief structures of the Flat Earth Society, but is nevertheless completely separate from the truth of birds migrating over a round planet.

2.4.4 Validation in Scientific Models: Kuppers and Lenhard Case Study

Kuppers and Lenhard's evaluation of validation of simulation in the natural and social sciences sought to demonstrate the relation between theory and validation in models (Günter et al. 2005). As a case study, they focused upon the infamous scenario in climate modelling of 'Arakawa's trick.'

Norman Phillis' model of atmospheric dynamics (Phillips 1956) was a very ambitious step toward climate modelling on a grand scale. The results of his simulation were viewed with respect by the research community of the time, demonstrating as they did a direct correspondence to empirically-observed patterns of airflow in the atmosphere, but his results were hampered by an unfortunate consequence of the equations: numerical instability prevented the model from making any long-term predictions.

Arakawa sought to solve this problem, and eventually did so by virtue of his notable trick: he altered the equations of state, incorporating assumptions that were not derived from conventional atmospheric theory, and in the process ensured long-term stability in the model. Understandably this technique was met with substantial skepticism, but eventually was accepted as further empirical data showed the accuracy of Arakawa's revised model despite the theoretical inadequacies.

Kuppers and Lenhard use this as a demonstration that 'performance beats theoretical accuracy' (Günter et al. 2005). In other words, a simulation can provide successful data without having a completely accurate representation of the phenomenon at hand. This certainly seems to spell trouble for Schmid's descriptions of relating truth in simulation to the theoretical background of that simulation. If we agree that simulations may achieve empirical accuracy despite theoretical inaccuracy, how does this affect our view of truth in simulation?

A different approach to the relation between model and theory seems required. The status of a successful, validated model relies on more than a simple correspondence between the model and the research community, or the model and its surrounding theoretical assumptions, as Arakawa's computational gambit demonstrates.

2.5 The Connection Between Theory and Simulation

2.5.1 Simulation as 'Miniature Theories'

The relations described thus far between theory and simulation clearly lack important elements. For example, while the coherence theory of truth is appealing in that, say, an evolutionary model may find validation by fitting cohesively into the existing theory of biological evolution, how that model relates to theory itself remains an open question. Likewise, if a model must fit into an existing set of beliefs, are we suddenly restricting the ability of simulation to generate new theory? Is there room for innovation in such concepts of validation?

One means to escape from this difficult connection between simulation and theory is to reform our definition completely: we may consider a simulation as a theory in itself. Within the simulation community this view is not uncommon:

> The validation problem in simulation is an explicit recognition that simulation models are like miniature scientific theories. Each of them is a set of propositions about how a particular manufacturing or service system works. As such, the warrant we give for these models can be discussed in the same terms that we use in scientific theorizing in general. (Kleindorfer et al. 1998, p. 1087)

Similarly, Colburn describes simulation as a means to 'test a hypothesis in a computer model of reality' (Colburn 2000, p. 172). In the context of Arakawa's trick, this perspective is attractive: in this view Arakawa's model can serve as a miniature theory of its own, and the perceived disconnect between the assumptions of his model and the accepted atmospheric theory are of no consequence to the validity of the model.

2.5.2 Simulations as Theory and Popperian Falsificationism

If we accept that simulations can take the form of such 'miniature theories,' then perhaps the question of validation becomes instead a question of the validity of scientific theories. Herskovitz suggests that the process of validating simulation models is at its root a Popperian process of falsification (Herskovitz 1991). In essence, given that a simulation model is considered validated if its results correspond to the behaviour of real-world systems, then a system must likewise be falsified if its results do not correspond to the real-world system.

However, Arakawa's trick once again throws this view into question. Arakawa's model by its very nature incorporates assumptions that are contrary to physical theory: as one example, he assumes that energy is conserved in his model of the atmosphere, while the real atmosphere does not display such conservation (Günter et al. 2005; Arakawa 1966). In this respect, is his model subject to Popperian falsification? If a central assumption of this model, one which informs every calculation of that model, is demonstrably false, does his model likewise lose all validity?

Kuppers and Lenhard argue that it does not: that the theory presented by Arakawa's model stands apart from the physical theory upon which it was initially based, and its performance speaks more to its validity than the accuracy of its conceptual assumptions. Likewise, we might imagine an evolutionary model falling victim to the same Popperian plight: assumptions made to simplify the process of evolution within the model may contradict observed facts about evolving species in nature. However, if those assumptions allow the model to display accuracy in another respect, either theoretical or empirical in nature, should we still decry the assumptions of that model?

2.5.3 The Quinean View of Science

In this context the Popperian view of falsification seems quite at odds with the potential nature of scientific models. In contrast to what Herskovitz seems to believe, simulations are far more than mere collections of assumptions designed to imitate and calculate the properties of natural phenomena. Indeed, simulations can often contain a rich backdrop of internal assumptions and theories, and the measure of a simulation's success seems likewise more rich than a simple comparison with accepted data and theory.

The Quinean view of science seems much more suited to the simulation endeavour (and indeed, many would argue, more suited to science of all varieties). The Duhem-Quine problem stands famously at odds with the Popperian view, asserting that scientific theories can in fact never be proved conclusively false on their own (Quine 1951, 1975). Given that theories depend on one or more (often many) related auxiliary assumptions, theories can be saved from definitive falsification by adjusting those auxiliary assumptions.

For example, Newton's laws of gravitation were able to explain a great deal of natural phenomena, and are still used extensively in modern physics. However, Newton's laws could not explain some clearly evident and anomalous behaviours in astronomical bodies: the perihelion of Mercury's orbit being a prime example. Yet, rather than simply disposing of Newton's theory as inadequate, scientists instead strove for another explanation in addition to Newton's theories, which later arrived in the guise of Einstein's general relativity. Now Newton's laws are presented as essentially a subset of Einstein's theory, displaying correct results within a certain set of reference frames. Likewise, points where Einstein's theories break down (i.e., at the Big Bang or at the event horizon of a singularity) are not taken as a falsification of Einstein's views, but rather an indication of the need for additional theories and assumptions to explain those anomalies in detail.

2.5.4 Simulation and the Quinean View

Having rightly discarded the Popperian view of simulation and embraced Quine's notion of flexible and interconnected scientific theories, we can revise our initial view of simulation in the context of this understanding. Noble (1998) provides a summary of one view of simulation in a Quinean context, specific to artificial life: he argues that the Quinean view implies that new models are generated according to a requirement to incorporate new information in an existing theory without completely reorganising that theory.

As an example Noble posits that a new simulation in artificial life may seek to explain a behaviour in biology as a consequence of an emergent process (Noble 1998). The modeler may then implement a model incorporating appropriate low-level assumptions with the intention of running the simulation to determine whether the expected behaviour does indeed emerge. In this respect the model is based upon pre-existing conceptual frameworks concerning the high-level behaviour, and contributes to the addition of this new behaviour into the overall biological theory by providing an explanation of that behaviour in terms of an emergent phenomenon.

More generally, we may add to this characterisation by referring back to the earlier discussion of simulation-as-theory. When constructing a model to allow for the integration of new information into an overall conceptual framework, a simulation model can function as an auxiliary hypothesis in and of itself: that simulation forms a theory, and thus is subject to the same standards as the larger conceptual framework. In this case, even if the model does not achieve validation in comparison to empirical data, all is not lost; in the appropriate Quinean fashion, the auxiliary hypotheses linked to that simulation may be revised to present a new version of the simulation-theory (perhaps by revising certain parameter values or similar).

Simulation then is not simply a means to simplify calculations within a pre-existing theoretical framework, it is a means to modify that theoretical framework. The validity of a simulation is not easy to determine by any means, but a simulation based in an existing framework that adds sensible assumptions to that framework may go a long way toward justifying its existence as a substantive part of the overall theory. Unlike in the Popperian view, an invalidated simulation need not be discarded, but instead revised; assumptions used in a simulation are pliable, and an alteration of same could allow that model to produce insights it originally appeared to lack.

2.6 ALife and Scientific Explanation

2.6.1 Explanation Through Emergence

Having established that simulation can perform a valuable role in the development of scientific theory, this analysis now turns to the role of simulation in scientific explanation specifically. The ability of simulation to provide complete and coherent scientific explanation will impact the strength with which this methodology can develop scientific theories; bearing this in mind, we require an understanding of the limits of simulation in developing explanations. This explanatory role for simulation is often hotly debated, particularly in the case of scientific models as described here (see the exchange between O'reilly and Farah 1999 and Burton and Young 1999 for one example, as the authors debate the explanatory coherence of distributed representations in psychological models).

Within ALife, which depends upon frequently abstract simulations of complex emergent systems, the explanation problem takes on a new dimension. As Noble posits (Noble 1998), ALife models provide the unique mechanism of emergence which can provide new elements of an explanation of a phenomenon; however, debate continues as to whether explanations of higher-order phenomena through emergence can capture a complete explanation of those phenomena. This debate is exemplified once more by the debate within cognitive science regarding connectionism and distributed representations: the reductive character of connectionist explanation of mental phenomena is seen as overly restrictive, removing the possibility of higher-order explanations of mental states.

As noted earlier in our discussion of connectionism, this debate is easily avoidable in one sense: if we accept that consciousness, for example, is a natural emergent property of neuronal activity, then the acceptance of this fact does not preclude the use of higher-order discussions of mental states as a means to explain the characteristics of that emergent consciousness. This does, however, seem to preclude the notion of that emergent explanation being a complete explanation; even if one can show that consciousness does indeed emerge from that lower-level activity, by the nature of emergent phenomena that consciousness is not easily reducible to those component activities.

2.6.2 Strong vs Weak Emergence

The variety of emergence discussed in the previous section is often referred to as 'strong emergence': the concept that not only is an emergent phenomenon difficult to reduce directly to its component parts, but in fact the emergent phenomenon can display downward causation, or supervenience (influencing its own component parts), thus making the cause of that emergent phenomenon very difficult to define. In essence, the whole is a consequence of its component parts, but is irreducible to the actions of those components (see O'Conner 1994; Nagel 1961).

Weak emergence, by contrast, is a means proposed by Mark Bedau to capture the emergent character of natural phenomena without the troublesome irreducibility (Bedau 1997). Bedau defines weak emergence thus:

> Macrostate P of S with microdynamic D is weakly emergent iff P can be derived from D and S 's external conditions but only by simulation. (Bedau 1997, p. 6)

Thus, similar to strong emergence, the macro-level behaviour of the emergent system cannot be predicted merely by knowledge of its micro-components. Crucially however, those macro-level properties can be derived by allowing those micro-components to perform their function in simulation. Under strong emergence, an evolutionary simulation constructed in bottom-up ALife fashion would be unable to capture the complete behaviour of the resultant emergent phenomenon. Under weak emergence, that simulation could indeed provide a derivation of that higher-level behaviour.

Bedau takes great pains to point out however that such weakly emergent behaviours are still, much like strongly emergent behaviours, essentially autonomous from their micro-level components. While in theory one could predict exactly the behaviour of a weakly emergent system with a perfectly accurate simulation of its micro-level components, in practice such simulations will be impossible to achieve. Instead, the micro-level explanation via simulation provides a means to observe the general properties of the macro-level weakly emergent result.

In this sense there appears to be a certain circularity to weak emergence: simulation can provide a micro-level explanation of an empirical phenomena, but in practice 'we can formulate and investigate the basic principles of weak emergent phenomena only by empirically observing them at the macro-level' (Bedau (1997), p. 25). Some may argue in fact that this constitutes an explanation in only a weak sense: one could point to a simulation of this type and note that the given micro-components lead to the specified macro-behaviour, but the level of insight into that macro-behaviour is still fundamentally limited despite intimate knowledge of the behaviour of its components.

This objection becomes important once more in the context of social simulation and the concept of non-reductive individualism, which is explored in Chaps. 5, 6 and 7. While Bedau's concept of weak emergence is less metaphysically tricky than classical strong emergence, the difficulties that remain in explanation by simulation despite this new categorisation of phenomena still allow for criticism of the simulation approach. Such criticisms will inform our discussion of social simulation in the second section of this text as well as our discussion of Alife in the current section.

2.6.3 Simulation as Thought Experiment

Unsurprisingly these difficulties in using simulation for scientific explanation have generated much discussion within the research community. Di Paolo, Noble and Bullock approach this thorny issue by proposing that simulations are best viewed as opaque thought experiments (Di Paolo et al. 2000). This proposal draws upon Bedau's earlier description of ALife models, describing them as 'computational thought experiments.'

A traditional thought experiment in this view constitutes 'in itself an explanation of its own conclusion and its implication' (Di Paolo et al. 2000, p. 6). In other words a thought experiment provides a self-contained means with which to probe the boundaries of the theory which informs that experiment. A successful thought experiment can provoke a reorganisation of an existing theory as it brings previously-known elements of that theory into a novel focus.

Simulation experiments, it is argued, can fulfill a similar purpose. However, simulations suffer from an inherent opacity: as noted in our discussion of emergence, the modeler's knowledge of the workings of the simulation do not imply an understanding of the simulation's results. Unlike in a conventional thought experiment, the modeler must spend time unraveling the result of his simulation, probing the consequences to determine the implications for theory.

As a result of this view, the authors propose a different methodology in simulation research than the conventional view. Firstly, they contend that the successful replication of a result given some mechanism described in the simulation does not constitute an explanation (a misconception common to simulation work, and clearly debunked by the characteristics of emergence mentioned earlier). In consequence the explanation which may be drawn from simulation work is likely to incorporate an 'explanatory organisation,' in which some elements of the problem are explained through micro-level behaviour, others may be explained at the macro-level, and still others in relations between the two.

In essence, they advocate an additional step in the modelling process in which the modeler performs experiments on the simulation, as one might do in a laboratory environment. The systematic exploration of the model itself is intended to provide a greater understanding of its inner workings, and in turn this theory of the model's behaviour must then be related to theories about the natural world which provide the inspiration for the model. So the ALife researcher can accept the view that emergent behaviours are difficult to explain via simulation, but at the same time forming theories about the simulation that relate to theories about those behaviours may produce a new insight into pre-existing theories, as with a successful thought experiment.

2.6.4 Explanation Compared: Simulations vs Mathematical Models

Taking the thought-experiment perspective into more depth, Bryden and Noble (2006) contrast the explanatory capacity of simulation models with that of mathematical models. They seek to explore what is required of an explanation derived from simulation, noting that the unfortunately commonly accepted view that a simple qualitative similarity between the simulation result and the behaviour of the real system is sufficient to provide any sort of explanation.

Bryden and Noble note another element of the inherent opacity of simulation research: the analytical incompleteness of such models. Mathematical treatments, when flawed, are easily revealed as such. Simulations in contrast can be run many times, with different parameter values, and flaws in the coding of the simulation may not be immediately apparent. Similarly, those simulation runs represent only isolated data points in the entire possible space of runs allowable in that model; since no researcher can spare the time to browse the entire space of parameter values for a simulation, the results we see are only a fraction of what is possible.

The authors go on to advocate a means for decomposing a simulation system into component mechanistic subsystems which are more amenable to mathematical explanation. A working model which is decomposed in this way may still not provide the complete analytical package of an exclusively mathematical treatment, but it is argued that this brings the researcher closer to a full analytical solution of the target system. Thus the computational model is seen as a means to generate the tools necessary to reach a cohesive mathematical explanation of the phenomenon under study.

In a broader context this approach is quite close to that proposed by Di Paolo, Noble and Bullock (Di Paolo et al. 2000). In both cases the modeler spends time exploring the confines of the model in question, probing its inner workings to define the parameters in which that model operates. Having done this, the modeler may begin to relate those theories about the model to theories about the world; in Bryden and Noble's view, for maximum explanatory power those relations should take the form of mathematical treatments. From both however we draw the conclusion that models which are able to replicate an emergent behaviour through the simulation of a system's micro-level component interactions is still very far from providing an explanation of that system. The simulation itself must be further deconstructed, its parameters understood, and its workings probed in order to relate that model successfully to the real system to which it relates. Thus simulation becomes a valuable tool in the scientist's repertoire, but one that must be supplemented by more traditional means of enquiry as well.

2.7 Summary and Conclusions

The problems facing the simulation methodology are clearly far from philosophically transparent. From mathematical models to evolutionary simulations, these tools display potential explanatory power and are quick to develop in comparison to traditional empirical studies. However, with that relative ease of use comes a correspondingly high difficulty of analysis.

The type of simulation discussed here, particularly in the context of ALife, focuses on the investigation of emergent phenomena in the natural world. These phenomena by their very nature are difficult to explain; simulation provides a means to view the origin of these phenomena from the behaviour of low-level component parts, but still lacks in its ability to explain the overall behaviour of that higher-level emergent order.

A number of researchers and philosophers have attempted to justify the use of simulation as a means for scientific explanation in a variety of ways; our synthesis up to this point indicates that simulation is certainly a useful element in the explanatory toolbox. Simulation however cannot stand alone: a simulation which displays an emergent behaviour still requires a theoretical framework to describe that behaviour.

Within ALife, different streams of thought have developed in response to these difficulties; questions surrounding the appropriate use of simulation in the field have led to extensive debate on the validity of simulation as a means to generate data. A further investigation into the theoretical underpinnings of ALife in the following chapter will provide insight into the fundamental aspects of this debate, and lead us further into this analysis of simulation methodology. This investigation will also provide important theoretical background for future discussion of Alife models and their relation to more general methodological concerns in modelling amongst the broader biology community.

References

Arakawa, A. (1966). Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. *Journal of Computational Physics, 1*, 119–143.

Bedau, M. A. (1997). Weak emergence. In J. Tomberlin (Ed.), *Philosophical perspectives: Mind, causation, and world* (Vol. 11, pp. 375–399). Oxford: Blackwell.

Bedau, M. A. (2003). Artificial life: Organization, adaptation and complexity from the bottom up. *Trends in Cognitive Sciences, 7*, 505–512.

Brooks, R. A. (1991). Intelligence without reason. In J. Myopoulos, & R. Reiter (Eds.), *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence* (pp. 569–595). San Mateo, CA: Morgan Kaufmann.

Brugmann, B., Tichy, W., & Jansen, N. (2004). Numerical simulation of orbiting black holes. *Physical Review Letters, 92*, 211101.1–211101.4.

Bryden, J., & Noble, J. (2006). Computational modelling, explicit mathematical treatments, and scientific explanation. In L. M. Rocha, et al. (Ed.), *Artificial Life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems* (pp. 520–526). Cambridge: MIT Press.

Buss, D. M. (2004). *Evolutionary psychology: The new science of the mind*. Boston, MA: Pearson.

Colburn, T. R. (2000). *Philosophy and computer science*. Armonk, NY: M.E. Sharpe.

Cox, T. J., & Loeb, A. (2007). The collision between the Milky Way and Andromeda. *Monthly Notices of the Royal Astronomical Society, 386*(1), 461–474.

Darwin, C. (1859). *The origin of species by means of natural selection*. London: John Murray.

Dawkins, R. (1995). *River out of Eden: A Darwinian view of life*. New York: Basic Books.

Di Paolo, E. A., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. In M. A. Bedau, et al. (Ed.), *Artificial life VII* (pp. 497–506). Cambridge, MA: MIT Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture. *Cognition, 28*, 3–71.

Fogel, D. B. (1988). An evolutionary approach to the travelling salesman problem. *Biological Cybernetics, 60*(2), 139–144.

Fraser, A. S. (1957). Simulation of genetic systems by automatic digital computers. *Australian Journal of Biological Science, 10*, 492–499.

Froese, T., Virgo, N., & Ikegami, T. (2014). Motility at the origin of life: Its characterization and a model. *Artificial Life, 20*(1), 55–76.

Goldberg, D. E. (1989). *Genetic algorithms in search*. Reading: Addison Wesley.

Gregory, J. M., et al. (2005). A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric Co2 concentration. *Geophysical Research Letters, 32*, L12703.1–L12703.5.

Gribbin, J. (1992). *Unveiling the edge of time*. New York: Harmony Books.

Herskovitz, P. J. (1991). A theoretical framework for simulation validation: Popper's falsificationism. *International Journal of Modelling and Simulation, 11*, 56–58.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press.

Humphreys, G. W., Freeman, G. W., & Muller, H. J. (1992). Lesioning a connectionist model of visual search: Selective effects on distractor grouping. *Canadian Journal of Psychology, 46*, 417–460.

Kleindorfer, G. B., O'Neill, L., & Ganeshan, R. (1998). Validation in simulation: Various positions in the philosophy of science. *Management Science, 44*(8), 1087–1099.

Küppers, G., & Lenhard, J. (2005). Validation of simulation: Patterns in the social and natural sciences. *Journal of Artificial Societies and Social Simulation, 8*(4), 3.

Langton, C. (2006). Artificial life. In C. Langton (Ed.), *Artificial life* (pp. 1–47). Reading: Addison-Wesley.

Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. In *CNLS'89: Proceedings of the Ninth Annual International Conference of the Center for Nonlinear Studies on Self-Organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks on Emergent Computation* (pp. 12–37). North-Holland Publishing Co.

Langton, C. G. (1992). Life at the edge of chaos. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *Artificial life II* (pp. 41–91). Reading: Addison-Wesley.

Law, A. M., & Kelton, W. D. (2000). *Simulation modeling and analysis third edition*. New York: McGraw Hill.

Luisi, P. L. (1998). About various definitions of life. *Origins of life and the evolution of the biosphere, 28*, 613–622.

Maher, J. (2007). Reconfigurable hardware evolution platform for a spiking neural-network controller. In *Reconfigurable Computing: Architectures, Tools and Applications, Third International Workshop, ARC 2007, Mangaratiba, 27–29 Mar 2007* (Lecture notes in computer science, Vol. 4419). Springer.

Matsuno, H., Hanczyc, M. M., & Ikegami, T. (2007). Self-maintained movements of droplets with convection flow. In M. Randall, H. A. Abbass, & J. Wiles (Eds.), *Progress in artificial life. ACAL 2007* (Lecture notes in computer science, Vol. 4828). Berlin/Heidelberg: Springer.

Maturana, H., & Varela, F. (1973). Autopoiesis and cognition: The realization of the living. In R. S. Cohen, & M. W. Wartofsky (Eds.), *Boston studies in the philosophy of science* (Vol. 42). Dordrecht: Springer.

Mitchell, M. (1996). *An introduction to genetic algorithms*. Cambridge: MIT Press.
Nagel, E. (1961). *The structure of science*. San Diego: Harcourt, Brace and Company.
Newell, A., & Simon, H. A. (1976). Computer science as empirical enquiry: Symbols and search. *Communications of the Association for Computing Machinery, 19*, 113–126.
Noble, J. (1998). *The evolution of animal communication systems: Questions of functions examined through simulation*. PhD Thesis, University of Sussex, Brighton.
O'Conner, T. (1994). Emergent properties. *American Philosophical Quarterly, 31*, 91–104.
Phillips, N. (1956). The general circulation of the atmosphere: A numerical experiment. *Quarterly Journal of the Royal Meteorological Society, 82*, 123–164.
Pinker, S., & Mehler, J. (1988). *Connections and symbols*. Cambridge, MA: MIT Press.
Quine, W. V. O. (1951). Two dogmas of empiricism. *Philosophical Review, 60*(1), 20–43.
Quine, W. V. O. (1975). On empirically equivalent systems of the world. *Erkenntnis, 9*, 313–328.
Rowlands, M. (1994). Connectionism and the language of thought. *British Journal for the Philosophy of Science, 45*(2), 485–503.
Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing*. Cambridge: MIT Press.
Sargent, R. G. (1982). Verification and validation of simulation models. In F. E. Cellier (Ed.), *Progress in modeling and simulation* (pp. 159–169). London: Academic Press.
Sargent, R. G. (1985). An expository on verification and validation of simulation models. In *Proceedings of the 17th Conference on Winter Simulation* (pp. 15–22). New York, NY: ACM Press.
Sasahara, K., & Ikegami, T. (2004). Song grammars as complex sexual displays. In J. Pollack, M. Bedau, P. Husbands, T. Ikegami, & R.A. Watson (Eds.), *Artificial life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems* (pp. 194–199). Cambridge, MA: MIT Press.
Sawyer, R. K. (2002). Nonreductive individualism: Part i: Supervenience and wild disjunction. *Philosophy of the Social Sciences, 32*, 537–561.
Sawyer, R. K. (2003). Artificial societies: Multi-agent systems and the micro-macro link in sociological theory. *Sociological Methods and Research, 31*(3), 37–75.
Sawyer, R. K. (2004). Social explanation and computational simulation. *Philosophical Explorations, 7*(3), 219–231.
Schlesinger, S., et al. (1979). Terminology for model credibility. *SCS Technical Communications for Modelling and Credibility in Simulation, 32*(3), 103–104.
Schmid, A. (2005). What is the truth of simulation? *Journal of Artificial Societies and Social Simulation, 8*(4), 5.
Von Neumann, J., & Burks, A. W. (1966). *Theory of self-reproducing automata*. Urbana: University of Illinois Press.